

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE EDUCACIÓN**  
**DEPARTAMENTO DE MÉTODOS DE INVESTIGACIÓN Y DIAGNÓSTICO EN**  
**EDUCACIÓN**



**TESIS DOCTORAL**

**Aprendizaje cooperativo e individual en el rendimiento académico en  
estudiantes universitarios:  
un meta-análisis**

**MEMORIA PARA OPTAR AL GRADO DE DOCTORA  
PRESENTADA POR**

**Celia Camilli Trujillo**

Directores

Eduardo López López  
María Castro Morera

**Madrid, 2015**

**Aprendizaje cooperativo e individual en el rendimiento  
académico en estudiantes universitarios:  
un meta-análisis**

Tesis doctoral realizada por  
**Celia Camilli Trujillo**

Bajo la dirección de  
Dr. Eduardo López López  
Dra. María Castro Morera



UNIVERSIDAD COMPLUTENSE DE MADRID

DEPARTAMENTO DE MÉTODOS DE INVESTIGACIÓN Y DIAGNÓSTICO EN  
EDUCACIÓN

Madrid, 2015





A Eduardo López, quien confió en mí sin conocerme. Gracias Eduardo

A María Castro, por adoptarme en mi orfandad y saber llevarme a término

A mis compañeros de vicisitudes por tantos momentos de contraste de nuestros avances

A mi hermano Alejandro, siempre estarás entre nosotros

A mi padre, modelo a seguir quien me ha enseñado que *never give up*

A mi madre, sus degustaciones culinarias han teñido cada una de estas páginas

A mi abuela, por su cercanía en mis horas de estudio

Y finalmente, a mi hija Daniela y a mi marido Max. A Daniela porque quiero que vea que no hay límites, que siempre se puede y que el esfuerzo rinde sus frutos. A Max por ser mi compañero y confidente, por estar a mi lado todo este tiempo ocupándote de mí

Max T.A.C.D.Y.L.



*A la tesis, la mía, la que hice al lado de mi familia*

Si alguien ha vivido una tesis, o ha vivido con un tesista las palabras que están en esa casa terminan casi siendo monosílabos. Sí, no, ahora sí, ahora no, bueno, por favor, gracias. ¿La leíste? ¿Tú qué crees? El comedor se convierte en estudio, en sala de investigación. La lámpara que está sobre la mesa no deja de estar encendida y el *wi-fi* no cesa en su titilar de conexión constante.

Los días, las semanas, los meses, los años pasan, las mismas palabras son las que se oyen, mientras se teclea, se buscan documentos, se procesan datos, se guardan. Las fundas con documentos crecen y el espacio se hace pequeño, muy pequeño. Las horas se suceden unas con otras, los días deben terminar en el sueño que no se sabe si repara o si más bien inquieta. Tal vez sea mejor no dormir y seguir, aunque no se sabe ni cómo.

Ya no hay baldas de almacenaje. Se inunda la parte de arriba de todos los armarios, de la biblioteca, se habilita un espacio en el trastero, junto al árbol de Navidad, justo entre las sillas de playa y el belén. Un poco más allá, la caja con los bañadores que esperan estar nuevamente de moda.

Cada vez que se celebran las uvas, con las doce campanadas, aunque sepamos que no se acaba ese año, nos prometemos que el próximo no habrá tesis. Es una promesa queda, silente, anhelante de poder abrazar sin tener pendientes las normas APA, las citas bien articuladas y la escritura del capítulo de discusión y conclusiones justo cuando las fuerzas están menguadas. La sensación es ambigua. ¡Si! ¡Que la tesis se termine! ¿Que se termine? ¡En serio! ¿En serio?

Las reuniones familiares son la oportunidad perfecta para desarrollar el tema favorito: meta-análisis. Las palabras claves ya son parte de la conversación. Saltan de la hoja de papel a la mesa bien servida para que mis padres, tíos, primos y demás allegados sepan de qué va la cosa, mientras un guiño cansado se cruza entre mi marido e hija. La terminología se apropia de todos y todos nos apropiamos de ella ¡Hasta la abuela nos observa con sorpresa! Nos ve desde lejos cómo el portátil entra y sale entre el postre y el café. Ve con especial interés cómo no levantamos la mirada de la pantalla y asiente con la cabeza, con la sabiduría de saber acompañar aunque las palabras parpadeantes se pierdan en su memoria.

Si se trata de un partido de fútbol de final de copa, el entretiempo, ese para reponer las palomitas, ir al baño y llenar los vasos es ocasión propicia para asomarse y ver cómo va el partido. Bueno, literalmente asomarse porque aunque trabajas con la televisión de fondo porque te permites ese día especial combinar trabajo con disfrute, una narrativa coherente de lo que ha sucedido es imposible. Días más tarde, cuando finalmente puedes conectarte con el *whatsapp* del grupo familiar es cuando te enteras luego de 57 mensajes sin leer lo que ha ocurrido. Te detienes por varios segundos frente al móvil pero decides no escribir porque “estás fuera de tiempo”.

¿Y la tesis? ¿En qué momento? ¿Cuándo? ¿Y ahora? Otra semana que se va sin dedicarle todo el tiempo que necesitaba. La imagen tan repetida de hojas de un almanaque que van cayendo al vacío se reproduce en tu cabeza. Los días no alcanzan.

De chicle. De chicle deberían ser los días.

Más fundas, más impresiones de las diferentes versiones. Se suceden una tras otra. Ya soy amiga del señor de la casa de impresiones. Sabe lo que quiero. No le digo nada, me presento con mis ojeras, esas que me acompañan tras los cristales de mis gafas, le dejo el pendrive y me imprime mientras me paseo por las agendas y las carpetas de colores... Voy a grabar un respaldo del respaldo porque no sé si este ordenador me está guardando todo. Yo creo que no. No confío.

La ciclo génesis explosiva -así lo hemos visto que lo llaman los meteorólogos en la televisión- podría adaptarse muy bien a esto de la tesis. Se definiría como aquel fenómeno que ocurre cuando se unen dos frentes: tesis y colegas. Resultado= ciclo-génesis-explosiva-compulsiva. El yo en el nosotros y el nosotros en el yo, el contar y el contarnos, el escuchar y el escucharnos se convierten en una catarsis colectiva. La pregunta matutina se repite todos los lunes ¿Avanzaste el fin de semana? El café con los del trabajo se va colmado de azúcar pero también de extra de cafeína. En las conversaciones en los pasillos no falta la típica y tan anhelada pregunta que en muchas ocasiones resulta tan lejana e inalcanzable que roza casi en lo esotérico ¿Cuándo defiendes? Un capuchino no vendría nada mal con buena nata y toque de canela.

Días, semanas, meses, años. El hábito se crea. La rutina se necesita. La voluntad se forma. La mente y el cuerpo ya conocen lo que se espera. Constancia, perseverancia y tenacidad. El pensamiento inunda el espacio. Conversaciones en voz alta en un monólogo sin público. Es un ir y venir permanente. Las páginas se convierten en una

catarata de letras articuladas compartidas con otros a quienes no he conocido. Confrontación con uno mismo. Anotaciones al margen. Borrar. Borrar una y otra vez. Aprender a releer, reescribir, reestructurar, repensar, recrear, redimensionar. El prefijo “re” se vuelve una constante.

Y así, cuando la mirada ve hacia atrás el camino recorrido, se abre un álbum de cromos: libros llenos de *post-it* que dejan caer vestigios del mar de los veranos; un sofá que ha acompañado lecturas hasta el amanecer; el sol que cobija mi ventana; la tierra, bendita tierra que me sostiene; la sapiencia de quienes me han guiado y el regazo del hogar, sentido y cómplice... son estampas cargadas de ganancias que curan las grietas de la calzada, que me llevan a dibujar en una nueva página el próximo proyecto, ese, que ya ha empezado a escribirse para cuando estas páginas se defiendan.



# ÍNDICE

<b>ABSTRACT</b>	<b>XIII</b>
<b>INTRODUCCIÓN</b>	<b>XXIII</b>
<b>PARTE TEÓRICA</b>	
<b>CAPÍTULO I. APRENDIZAJE COOPERATIVO: FUNDAMENTOS TEÓRICOS</b>	<b>1</b>
<b>1.1. TEORÍAS PSICOPEDAGÓGICAS DEL APRENDIZAJE COOPERATIVO</b>	<b>1</b>
1.1.1. TEORÍAS DEL APRENDIZAJE CONDUCTUAL	1
1.1.2. PIAGET Y EL CONFLICTO COGNITIVO	3
1.1.3. LOS NEOPIAGETIANOS Y EL CONFLICTO SOCIOCOGNITIVO	5
1.1.4. VIGOTSKY Y EL EXPERTO EN LA INTERACCIÓN SOCIAL	6
1.1.5. TEORÍA DE LA INTERDEPENDENCIA SOCIAL	8
1.1.5. TEORÍA DE LA CONTROVERSA CONSTRUCTIVA	11
<b>1.2. APRENDIZAJE INDIVIDUAL, COMPETITIVO Y COOPERATIVO</b>	<b>12</b>
<b>1.3. APRENDIZAJE COOPERATIVO Y APRENDIZAJE COLABORATIVO</b>	<b>18</b>
<b>1.4. ELEMENTOS ESENCIALES DEL APRENDIZAJE COOPERATIVO</b>	<b>23</b>
<b>1.5. TIPOS DE GRUPOS DE APRENDIZAJE COOPERATIVO</b>	<b>29</b>
<b>1.6. VARIABLES MODERADORAS QUE INCIDEN EN EL APRENDIZAJE COOPERATIVO</b>	<b>32</b>
1.6.1 VARIABLES COGNITIVAS	32
1.6.2. VARIABLES SOCIALES	34
1.6.3. VARIABLES INSTRUCCIONALES	35
<b>1.7. MÉTODOS COOPERATIVOS</b>	<b>41</b>
1.7.1. <i>STUDENT TEAM LEARNING (STL)</i>	45
1.7.1.1. <i>Student Teams–Achievement Divisions (STAD)</i>	46
1.7.1.2. <i>Teams–Games–Tournament (TGT)</i>	47
1.7.1.3. <i>Jigsaw II</i>	48
1.7.2. <i>COOPERATIVE INVESTIGATION</i>	49
1.7.2.1. <i>Group Investigation (GI)</i>	50
1.7.2.2. <i>Jigsaw I</i>	50
1.7.3. <i>JIGSAW Y SUS VARIANTES</i>	51
1.7.4. <i>DYADIC METHODS</i>	52
1.7.4.1. <i>Peer Collaboration</i>	54
1.7.4.2. <i>Reciprocal Peer Tutoring (RPT)</i>	55
1.7.4.3. <i>Scripted Cooperation (SC)</i>	56
1.7.4.4. <i>Peer Teaching</i>	57
1.7.4.5. <i>Think–Pair–Share</i>	57
1.7.5. <i>LEARNING TOGETHER (LT)</i>	58
1.7.6. <i>INFORMAL METHODS</i>	59
1.7.7. <i>COOPERATIVE LEARNING AND COMPUTERS</i>	60
1.7.8. LOS MÉTODOS COOPERATIVOS Y SUS DIFERENCIAS	63
<b>CAPÍTULO II. APRENDIZAJE COOPERATIVO: EVIDENCIAS EMPÍRICAS</b>	<b>69</b>
<b>2.1. RESULTADOS DEL APRENDIZAJE COOPERATIVO</b>	<b>69</b>
2.1.1. PERSPECTIVA INTERNACIONAL DE INVESTIGACIONES EN LA UNIVERSIDAD	70
2.1.2. PERSPECTIVA ESPAÑOLA DE INVESTIGACIONES EN LA UNIVERSIDAD	78
<b>2.2. EFICACIA DEL APRENDIZAJE COOPERATIVO EN EL RENDIMIENTO: META–ANÁLISIS</b>	<b>82</b>
2.2.1. EVIDENCIA BASADA EN INVESTIGACIÓN EDUCATIVA	84
2.2.2. META–ANÁLISIS EN LOS QUE SE INCLUYE A LA UNIVERSIDAD	88
2.2.3. META–ANÁLISIS CENTRADOS SOLO EN LA UNIVERSIDAD	98



## **PARTE EMPÍRICA**

### **CAPÍTULO III. JUSTIFICACIÓN Y OBJETIVOS DEL ESTUDIO 109**

### **CAPÍTULO IV. MÉTODO 117**

#### **4.1. REVISIÓN DE LA LITERATURA 117**

#### **4.2. CRITERIOS DE INCLUSIÓN/EXCLUSIÓN 121**

##### 4.2.1. FUENTES PRIMARIAS 122

##### 4.2.2. FUENTES SECUNDARIAS 124

#### **4.3. CODIFICACIÓN DE LOS ESTUDIOS 128**

##### 4.3.1. CARACTERÍSTICA DE LOS ESTUDIOS 128

##### 4.3.2. MANUAL DE CODIFICACIÓN 132

#### **4.4. CALIDAD METODOLÓGICA DE LOS ESTUDIOS: LA ESCALA 132**

##### 4.4.1. PRIMERA FASE: CONSTRUCCIÓN Y DISEÑO 133

##### 4.4.2. SEGUNDA FASE: JUICIO DE EXPERTOS 136

###### 4.4.2.1. Los expertos 137

###### 4.4.2.2. Consistencia interna 140

###### 4.4.2.3. Comentarios y sugerencias generales 141

###### 4.4.2.4. Las dimensiones y los ítems 142

###### 4.4.2.5. La puntuación global 153

##### 4.4.3. TERCERA FASE: FIABILIDAD INTER-OBSERVADORES 155

#### **4.5. MEDIDA DE LOS RESULTADOS 158**

##### 4.5.1. PROGRAMA Y SOFTWARE ESTADÍSTICO 158

##### 4.5.2. UNIDAD DE ANÁLISIS ESTADÍSTICO 159

##### 4.5.3. MAGNITUD DEL EFECTO GLOBAL 164

##### 4.5.4. ESTIMACIÓN GLOBAL DEL EFECTO 167

##### 4.5.5. HETEROGENEIDAD DE LOS ESTUDIOS 168

##### 4.5.6. VARIABLES MODERADORAS 170

##### 4.5.7. ANÁLISIS DE SENSIBILIDAD 171

##### 4.5.8. SESGO DE PUBLICACIÓN 172

### **CAPÍTULO V. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS 175**

#### **5.1. RESULTADOS DEL ANÁLISIS 175**

##### 5.1.1. DESCRIPCIÓN DE LA MUESTRA 175

###### 5.1.1.1. Variables extrínsecas 176

###### 5.1.1.2. Variables metodológicas 187

###### 5.1.1.3. Variables ecológicas 191

##### 5.1.2. SÍNTESIS CUANTITATIVA 202

###### 5.1.2.1. Magnitud del efecto global: el rendimiento 202

###### 5.1.2.2. Análisis de sensibilidad 207

###### 5.1.2.3. Heterogeneidad 211

###### 5.1.2.4. Análisis por subgrupos 212

###### 5.1.2.4.1. Variables extrínsecas 216

###### 5.1.2.4.2. Variables metodológicas 219

###### 5.1.2.4.3. Variables ecológicas 223

###### 5.1.2.5. Análisis por estratos 235

###### 5.1.2.6. Sesgo de publicación 242

##### 5.1.3. SÍNTESIS CUALITATIVA: CALIDAD DE LOS ESTUDIOS 244

#### **5.2. DISCUSIÓN DE RESULTADOS 251**

### **CAPÍTULO VI. CONCLUSIONES E IMPLICACIONES EDUCATIVAS 285**

### **CAPÍTULO VII. LIMITACIONES Y PROSPECTIVA 297**

<b>REFERENCIAS</b>	<b>311</b>
<b>ANEXOS</b>	<b>367</b>
ANEXO 1. GLOSARIO DE TÉRMINOS MÉTODOS COOPERATIVOS (INGLÉS–ESPAÑOL)	368
ANEXO 2. CRITERIOS DE CALIDAD EN ORGANIZACIONES AMERICANAS E INGLASAS	369
ANEXO 3. EVIDENCIAS BÚSQUEDA DE FUENTES PRIMARIAS	370
ANEXO 4. LISTADO ESTUDIOS DEL META–ANÁLISIS DE BOWEN 2000	378
ANEXO 5. LISTADO ESTUDIOS DEL META–ANÁLISIS DE SPRINGER, STANNE Y DONOVAN 1999	379
ANEXO 6. LISTADO ESTUDIOS DEL META–ANÁLISIS DE JOHNSON Y JOHNSON 1989	381
ANEXO 7. CLASIFICACIÓN INTERNACIONAL NORMALIZADA DE LA EDUCACIÓN (CINE)	388
ANEXO 8. MANUAL DE CODIFICACIÓN	390
ANEXO 9. VERSIÓN INICIAL INSTRUMENTO PARA MEDIR LA CALIDAD METODOLÓGICA DE LOS ESTUDIOS (VERSIÓN EN ESPAÑOL)	396
ANEXO 10. VERSIÓN INICIAL INSTRUMENTO PARA MEDIR LA CALIDAD METODOLÓGICA DE LOS ESTUDIOS (VERSIÓN EN INGLÉS)	398
ANEXO 11. SOLICITUD JUICIO DE EXPERTOS (VERSIÓN EN ESPAÑOL)	400
ANEXO 12. SOLICITUD JUICIO DE EXPERTOS (VERSIÓN EN INGLÉS)	406
ANEXO 13. RESULTADOS JUICIO DE EXPERTOS: FRECUENCIAS, PORCENTAJES Y COEFICIENTE ALFA CRONBACH	412
ANEXO 14. COMENTARIOS Y SUGERENCIAS GENERALES DE LOS EXPERTOS	414
ANEXO 15. PROCESO DE MODIFICACIÓN DE LAS DIMENSIONES E ÍTEMS DE LA ESCALA SEGÚN JUICIO DE EXPERTOS	417
ANEXO 16. VERSIÓN FINAL DE LA ESCALA DE CALIDAD METODOLÓGICA SEGÚN RECOMENDACIONES DEL JUICIO DE EXPERTOS	448
ANEXO 17. ESTUDIOS SELECCIONADOS AL AZAR PARA LA FIABILIDAD INTER–CODIFICADORES	456
ANEXO 18. RESULTADO ÍNDICE DE KAPPA	457
ANEXO 19. VERSIÓN FINAL DE LA ESCALA DE CALIDAD METODOLÓGICA SEGÚN FIABILIDAD INTER-OBSERVADORES	458
ANEXO 20. NÚMERO DE RESULTADOS POR ESTUDIOS PRIMARIOS	466
ANEXO 21. FOREST PLOT CALIDAD METODOLÓGICA ESTUDIOS PRIMARIOS	468
ANEXO 22. RESUMEN CARACTERÍSTICAS PRINCIPALES DE LOS ESTUDIOS INCLUIDOS EN EL META–ANÁLISIS	469

## ÍNDICE DE TABLAS

TABLA 1: DIFERENCIAS ENTRE APRENDIZAJE INDIVIDUAL, COOPERATIVO Y COMPETITIVO	16
TABLA 2: ROL DOCENTE Y DEL ALUMNO EN LOS MÉTODOS COOPERATIVOS	41
TABLA 3: CLASIFICACIÓN DE LOS MÉTODOS COOPERATIVOS SEGÚN SLAVIN	42
TABLA 4: CLASIFICACIÓN DE LOS MÉTODOS COOPERATIVOS SEGÚN KAGAN	43
TABLA 5: PRINCIPALES MÉTODOS COOPERATIVOS EN LA UNIVERSIDAD	45
TABLA 6: CUADRO COMPARATIVO <i>JIGSAW II</i> , <i>STAD</i> Y <i>TGT</i>	49
TABLA 7: COMPARATIVA ENTRE LOS MÉTODOS <i>JIGSAW I</i> , <i>II</i> , <i>III</i> Y <i>IV</i>	52
TABLA 8: DIFERENCIAS ENTRE <i>PEER TUTORING</i> Y <i>PEER COLLABORATION</i>	55
TABLA 9: CLASIFICACIÓN MÉTODOS COOPERATIVOS SEGÚN ESTRUCTURA DE RECOMPENSA	64
TABLA 10: CLASIFICACIÓN MÉTODOS COOPERATIVOS SEGÚN ESTRUCTURA DE RECOMPENSAS, TAREAS E IGUALDAD DE OPORTUNIDADES PARA LA PUNTUACIÓN	65
TABLA 11: RESULTADOS POR TIPO DE TAREA Y NIVEL EDUCATIVO DEL META-ANÁLISIS DE QIN ET AL. 1995	91
TABLA 12: RESULTADOS MÉTODOS COOPERATIVOS DEL META-ANÁLISIS DE JOHNSON JOHNSON & STANNE (2000)	97
TABLA 13: RESULTADOS RENDIMIENTO DEL META-ANÁLISIS DE SPRINGER, STANNE & DONOVAN (1999)	99
TABLA 14: RESULTADOS VARIABLES MODERADORAS DEL META-ANÁLISIS DE SPRINGER, STANNE & DONOVAN (1999)	100
TABLA 15: RESULTADOS VARIABLES MODERADORAS DEL META-ANÁLISIS DE SPRINGER, STANNE & DONOVAN (1999)	101
TABLA 16: RESULTADOS VARIABLES MODERADORAS DEL META-ANÁLISIS DE SPRINGER, STANNE & DONOVAN (1999)	102
TABLA 17: CARACTERÍSTICAS DE LOS ESTUDIOS DEL META-ANÁLISIS DE JOHNSON & JOHNSON (2002A)	104
TABLA 18: FUENTES SECUNDARIAS CONSULTADAS: META-ANÁLISIS	119
TABLA 19: FUENTES SECUNDARIAS CONSULTADAS: OTROS ESTUDIOS PRIMARIOS	120
TABLA 20: NÚMERO TOTAL DE ESTUDIOS PROVENIENTES DE LA REVISIÓN DE LA LITERATURA	122
TABLA 21: NÚMERO TOTAL DE ESTUDIOS EXCLUIDOS E INCLUIDOS DE LAS FUENTES PRIMARIAS	123
TABLA 22: NÚMERO TOTAL DE ESTUDIOS PROVENIENTES DE LAS FUENTES SECUNDARIAS	125
TABLA 23: NÚMERO TOTAL DE ESTUDIOS INCLUIDOS/EXCLUIDOS DE LA REVISIÓN DE META-ANÁLISIS	126
TABLA 24: NÚMERO TOTAL DE ESTUDIOS INCLUIDOS/EXCLUIDOS DE LA REVISIÓN DE OTROS ESTUDIOS PRIMARIOS	126
TABLA 25: NÚMERO TOTAL DE ESTUDIOS EXCLUIDOS/INCLUIDOS DE LA REVISIÓN NARRATIVA	127
TABLA 26: VERSIÓN INICIAL DEL INSTRUMENTO DE CALIDAD METODOLÓGICA	134
TABLA 27: COEFICIENTE ALPHA DE CRONBACH PARA CADA CRITERIO DE VALORACIÓN DEL INSTRUMENTO	140
TABLA 28: COEFICIENTE DE ALPHA DE CRONBACH POR CADA DIMENSIÓN DEL INSTRUMENTO	141
TABLA 29: VALORACIÓN DICOTÓMICA DE UN INDICADOR	142
TABLA 30: VALORACIÓN POLITÓMICA DE UN INDICADOR	142
TABLA 31: DIMENSIONES DE LA ESCALA ANTES Y DESPUÉS DEL JUICIO DE EXPERTOS	143
TABLA 32: VERSIÓN INICIAL Y FINAL ÍTEM 1.3	144

TABLA 33: VERSIÓN INICIAL Y FINAL DEL ÍTEM 2.1	145
TABLA 34: VERSIÓN INICIAL Y FINAL DEL ÍTEM 3.1	146
TABLA 35: VERSIÓN INICIAL Y FINAL DEL ÍTEM 3.2	147
TABLA 36: VERSIÓN INICIAL Y FINAL DEL ÍTEM 3.3	148
TABLA 37: VERSIÓN INICIAL Y FINAL DEL ÍTEM 6.1	149
TABLA 38: VERSIÓN INICIAL Y FINAL DEL ÍTEM 6.2	150
TABLA 39: NUEVOS ÍTEMS RELACIONADOS CON LA VALIDEZ DE CONSTRUCTO	151
TABLA 40: NUEVOS ÍTEMS RELACIONADOS CON LA VALIDEZ INTERNA	152
TABLA 41: NUEVOS ÍTEMS RELACIONADOS CON LA VALIDEZ INTERNA	152
TABLA 42: NUEVOS ÍTEMS RELACIONADOS CON LA VALIDEZ EXTERNA	152
TABLA 43: NUEVOS ÍTEMS VALIDEZ DE CONCLUSIÓN ESTADÍSTICA	153
TABLA 44: COMENTARIOS Y SUGERENCIAS VALORACIÓN PUNTUACIÓN GLOBAL	153
TABLA 45: VERSIÓN INICIAL Y FINAL DE LA PUNTUACIÓN GLOBAL DE LA ESCALA	155
TABLA 46: MODIFICACIONES DE LOS INDICADORES LUEGO DE LA FIABILIDAD INTER-OBSERVADORES	157
TABLA 47: TIPOS DE VALIDEZ Y PORCENTAJE ASIGNADO ANTES Y DESPUÉS DE LA FIABILIDAD INTER-OBSERVADORES	158
TABLA 48: DOCUMENTOS QUE REPORTAN MAS DE UN PROYECTO DE INVESTIGACIÓN O RESULTADO	160
TABLA 49: DOCUMENTOS QUE REPORTAN VARIOS GRUPOS EXPERIMENTALES Y UN GRUPO CONTROL	161
TABLA 50: DOCUMENTOS QUE REPORTAN VARIOS GRUPOS EXPERIMENTALES Y UN GRUPO CONTROL (CONTINUACIÓN)	161
TABLA 51: DOCUMENTOS QUE REPORTAN VARIOS RESULTADOS DE RENDIMIENTO	164
TABLA 52: NÚMERO TOTAL DE ESTUDIOS PROVENIENTES DE LA REVISIÓN DE LA LITERATURA	176
TABLA 53: MUESTRA FINAL	176
TABLA 54: RESUMEN CARACTERÍSTICAS EXTRÍNECAS DE LA MUESTRA	177
TABLA 55: LISTADO REVISTAS CIENTÍFICAS ESTUDIOS INCLUIDOS	182
TABLA 56: RELACIÓN N° DE AUTORES POR ESTUDIOS INCLUIDOS	183
TABLA 57: RESULTADOS PATRÓN DE COLABORACIÓN ENTRE AUTORES	184
TABLA 58: RESULTADOS N° DE INSTITUCIONES POR ESTUDIOS INCLUIDOS	185
TABLA 59: RESULTADOS PATRÓN DE COLABORACIÓN ENTRE INSTITUCIONES	186
TABLA 60: RESUMEN CARACTERÍSTICAS METODOLÓGICAS DE LA MUESTRA	188
TABLA 61: RESUMEN CARACTERÍSTICAS ECOLÓGICAS DE LA MUESTRA	191
TABLA 62: RESULTADOS MÉTODOS COOPERATIVOS CON Y SIN CLASIFICACIÓN ESTUDIOS INCLUIDOS	196
TABLA 63: MAGNITUD DEL EFECTO GLOBAL: MODELO EFECTOS FIJOS Y ALEATORIOS	202
TABLA 64: ANÁLISIS DE SENSIBILIDAD	208
TABLA 65: RESUMEN MAGNITUDES DEL EFECTO POR SUBGRUPOS SEGÚN VARIABLES EXTRÍNECAS, METODOLÓGICAS Y ECOLÓGICAS	213
TABLA 66: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: 1980-1999 Y 2000-2012	216
TABLA 67 RESULTADOS DEL ANOVA VARIABLE MODERADORA: 1980-1999 Y 2000-2012	217
TABLA 68: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: LUSTROS	217
TABLA 69: RESULTADOS ANOVA VARIABLE MODERADORA: LUSTROS	218
TABLA 70: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: DÉCADAS	218

TABLA 71: RESULTADOS ANOVA VARIABLE MODERADORA: DÉCADAS	218
TABLA 72: RESULTADOS ANOVA VARIABLE MODERADORA: TIPOS DE DOCUMENTO	219
TABLA 73: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: TIPOS DE DOCUMENTO	219
TABLA 74: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: ASIGNACIÓN	220
TABLA 75: RESULTADOS ANOVA VARIABLE MODERADORA: ASIGNACIÓN	220
TABLA 76: RESULTADOS DE LA MAGNITUD DEL EFECTO Y LAS PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: ASIGNACIÓN AZAR Y NO AZAR	221
TABLA 77: RESULTADOS ANOVA VARIABLE MODERADORA: ASIGNACIÓN AZAR Y NO AZAR	221
TABLA 78: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: SITUACIÓN BASAL	221
TABLA 79: RESULTADOS DEL ANOVA VARIABLE MODERADORA: SITUACIÓN BASAL	221
TABLA 80: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBA DE HOMOGENEIDAD VARIABLE MODERADORA: SEGSOS DEL INVESTIGADOR	222
TABLA 81: RESULTADOS DEL ANOVA VARIABLE MODERADORA: SEGSOS DEL INVESTIGADOR	222
TABLA 82: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: INSTRUMENTOS DE MEDIDA	223
TABLA 83: RESULTADOS ANOVA VARIABLE MODERADORA: INSTRUMENTOS DE MEDIDA	223
TABLA 84: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: ASIGNATURAS	225
TABLA 85: RESULTADOS DEL ANOVA VARIABLE MODERADA: ASIGNATURAS	225
TABLA 86: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: CLASIFICACIÓN DE LA EDUCACIÓN Y FORMACIÓN	226
TABLA 87: RESULTADOS DEL ANOVA VARIABLE MODERADORA: ÁREAS DE CONOCIMIENTO	227
TABLA 88: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: CLASIFICACIÓN MÉTODOS	228
TABLA 89: RESULTADOS DEL ANOVA VARIABLE MODERADORA: CLASIFICACIÓN MÉTODOS COOPERATIVOS	228
TABLA 90: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: DURACIÓN DEL PROGRAMA	229
TABLA 91: RESULTADOS DEL ANOVA VARIABLE MODERADORA: DURACIÓN DEL PROGRAMA	229
TABLA 92: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: CONFORMACIÓN DE LOS GRUPOS	230
TABLA 93: RESULTADOS DEL ANOVA VARIABLE MODERADORA: CONFORMACIÓN DE LOS GRUPOS	230
TABLA 94: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: N° DE INTEGRANTES	231
TABLA 95: RESULTADOS DEL ANOVA VARIABLE MODERADORA: N° DE INTEGRANTES	231
TABLA 96: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: GRUPOS PARES E IMPARES	232
TABLA 97: RESULTADOS DEL ANOVA VARIABLE MODERADORA: GRUPOS PARES E IMPARES	232
TABLA 98: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS DE HOMOGENEIDAD VARIABLE MODERADORA: ESTRUCTURA DE RECOMPENSAS	233
TABLA 99: RESULTADOS DEL ANOVA VARIABLE MODERADORA: ESTRUCTURA DE RECOMPENSAS	233

TABLA 100: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: ESTRUCTURA DE TAREA	234
TABLA 101: RESULTADOS ANOVA VARIABLE MODERADORA: ESTRUCTURA DE TAREA	234
TABLA 102: RESULTADOS DEL ANOVA VARIABLE MODERADORA: IGUALDAD OPORTUNIDADES PUNTUACIÓN	235
TABLA 103: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD VARIABLE MODERADORA: IGUALDAD OPORTUNIDADES PUNTUACIÓN	235
TABLA 104: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD ESTRATO: CIENCIAS + MÉTODOS COOPERATIVOS	237
TABLA 105: RESULTADOS ANOVA ESTRATO: CIENCIAS + MÉTODOS COOPERATIVOS	237
TABLA 106: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD ESTRATO: CIENCIAS SOCIALES + MÉTODOS COOPERATIVOS	239
TABLA 107: RESULTADOS ANOVA ESTRATO: CIENCIAS SOCIALES + MÉTODOS COOPERATIVOS	239
TABLA 108: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD ESTRATO: HUMANIDADES/ARTES + MÉTODOS COOPERATIVOS	241
TABLA 109: RESULTADOS ANOVA ESTRATO: HUMANIDADES/ARTES + MÉTODOS COOPERATIVOS	241
TABLA 110: RESULTADOS MAGNITUD DEL EFECTO Y PRUEBAS HOMOGENEIDAD ESTRATO: CALIDAD METODOLÓGICA	244
TABLA 111: RESULTADOS DEL ANOVA VARIABLE MODERADORA: CALIDAD METODOLÓGICA	245
TABLA 112: RESUMEN MAGNITUD DEL EFECTO ESTRATO: CALIDAD METODOLÓGICA + AÑOS DE PUBLICACIÓN + MÉTODOS COOPERATIVOS + ÁREAS DE CONOCIMIENTO	246
TABLA 113: RESULTADOS MAGNITUD DEL EFECTO ESTRATO: CALIDAD METODOLÓGICA + AÑOS DE PUBLICACIÓN	247
TABLA 114: RESULTADOS MAGNITUD DEL EFECTO ESTRATO: CALIDAD METODOLÓGICA + MÉTODOS COOPERATIVOS	249
TABLA 115: RESULTADOS MAGNITUD DEL EFECTO ESTRATO: CALIDAD METODOLÓGICA + ÁREAS DE CONOCIMIENTO	250
TABLA 116: RESUMEN RESULTADOS DEL ANOVA ESTRATOS CALIDAD METODOLÓGICA	251
TABLA 117: CONTRASTE ENTRE RESULTADOS OBTENIDOS (MAGNITUD DEL EFECTO)	Y
OTRAS SÍNTESIS CUANTITATIVAS	270

## ÍNDICE DE GRÁFICOS

GRÁFICO 1: RESULTADOS POR DÉCADAS ESTUDIOS INCLUIDOS	178
GRÁFICO 2: RESULTADOS POR LUSTROS ESTUDIOS INCLUIDOS	178
GRÁFICO 3: RESULTADOS POR TIPOS DE FUENTE ESTUDIOS INCLUIDOS	179
GRÁFICO 4: RESULTADOS POR TIPOS DE DOCUMENTO DE LOS ESTUDIOS INCLUIDOS	180
GRÁFICO 5: RESULTADOS LUGAR DE PROCEDENCIA ESTUDIOS INCLUIDOS	181
GRÁFICO 6: ASIGNACIÓN AL AZAR DE LOS SUJETOS	188
GRÁFICO 7: SITUACIÓN BASAL	189
GRÁFICO 8: SESGO DEL INVESTIGADOR	189
GRÁFICO 9: INSTRUMENTOS DE MEDIDA	190
GRÁFICO 10: CALIDAD METODOLÓGICA	191
GRÁFICO 11: RESULTADOS POR ASIGNATURAS ESTUDIOS INCLUIDOS	193
GRÁFICO 12: RESULTADOS POR CLASIFICACIÓN SEGÚN EDUCACIÓN Y FORMACIÓN ESTUDIOS INCLUIDOS	194
GRÁFICO 13: RESULTADOS MÉTODOS COOPERATIVOS ESTUDIOS INCLUIDOS	195
GRÁFICO 14: CLASIFICACIÓN DE MÉTODOS COOPERATIVOS INCLUIDOS	197
GRÁFICO 15: RESULTADOS N° DE INTEGRANTES ESTUDIOS INCLUIDOS	198
GRÁFICO 16: INTEGRANTES AGRUPADOS	199
GRÁFICO 17: CONFORMACIÓN DE GRUPOS	199
GRÁFICO 18: DURACIÓN DEL PROGRAMA	200
GRÁFICO 19: DIFERENCIAS ENTRE RECOMPENSA, TAREA Y PUNTUACIÓN	201

## ÍNDICE DE ILUSTRACIONES

ILUSTRACIÓN 1: CLASIFICACIÓN DE LOS MÉTODOS COOPERATIVOS SEGÚN SLAVIN Y KAGAN	44
ILUSTRACIÓN 2: RESULTADOS DEL APRENDIZAJE COOPERATIVO Y PATRONES DE INTERACCIÓN	69
ILUSTRACIÓN 3: DIAGRAMA DE FLUJO INCLUSIÓN/EXCLUSIÓN DE ESTUDIOS DE FUENTES PRIMARIAS	124
ILUSTRACIÓN 4: NÚMERO TOTAL DE ESTUDIOS EXCLUIDOS/INCLUIDOS DE LAS FUENTES PRIMARIAS	127
ILUSTRACIÓN 5: NÚMERO TOTAL DE ESTUDIOS INCLUIDOS/EXCLUIDOS DE LAS FUENTES SECUNDARIAS	128
ILUSTRACIÓN 6: RESULTADOS CLUSTER DE COLABORACIÓN ENTRE AUTORES	185
ILUSTRACIÓN 7: RESULTADOS CLUSTER DE COLABORACIÓN ENTRE INSTITUCIONES	187
ILUSTRACIÓN 8: <i>FLOREST PLOT</i> : MAGNITUD DEL EFECTO DE TODOS LOS ESTUDIOS INCLUIDOS	204
ILUSTRACIÓN 9: <i>FLOREST PLOT</i> : MAGNITUD DE EFECTOS GRANDES Y SIGNIFICATIVOS DE LOS ESTUDIOS INCLUIDOS	206
ILUSTRACIÓN 10: <i>FLOREST PLOT</i> : MAGNITUD DE EFECTOS NEGATIVOS DE LOS ESTUDIOS INCLUIDOS	207
ILUSTRACIÓN 11: <i>FLOREST PLOT</i> MAGNITUD DEL EFECTO ESTRATO: CIENCIAS + MÉTODOS COOPERATIVOS	238
ILUSTRACIÓN 12: <i>FLOREST PLOT</i> MAGNITUD DEL EFECTO ESTRATO: CIENCIAS SOCIALES + MÉTODOS COOPERATIVOS	240
ILUSTRACIÓN 13: <i>FLOREST PLOT</i> MAGNITUD DEL EFECTO ESTRATO: HUMANIDADES/ARTES + MÉTODOS COOPERATIVOS	242
ILUSTRACIÓN 14: <i>FUNNEL PLOT</i> : SESGO DE PUBLICACIÓN	242
ILUSTRACIÓN 15: GRÁFICO DE EGGER: SESGO DE PUBLICACIÓN	243
ILUSTRACIÓN 16 ELEMENTOS EFICACES PARA UN APRENDIZAJE COOPERATIVO EN LA UNIVERSIDAD (EFECTOS GRANDES)	274
ILUSTRACIÓN 17: ELEMENTOS EFICACES PARA UN APRENDIZAJE COOPERATIVO EN LA UNIVERSIDAD (EFECTOS MODERADOS)	275
ILUSTRACIÓN 18: LA CALIDAD METODOLÓGICA Y SU INCIDENCIA EN EL RENDIMIENTO ACADÉMICO EN LA UNIVERSIDAD	277
ILUSTRACIÓN 19: NIVELES DE CALIDAD METODOLÓGICA Y MAGNITUDES DEL EFECTO	278
ILUSTRACIÓN 20: ELEMENTOS EFICACES DEL APRENDIZAJE COOPERATIVO PARA LA ENSEÑANZA DE LAS CIENCIAS	279
ILUSTRACIÓN 21: ELEMENTOS EFICACES DEL APRENDIZAJE COOPERATIVO PARA LA ENSEÑANZA DE LAS CIENCIAS SOCIALES Y HUMANIDADES Y ARTES	282
ILUSTRACIÓN 22: ELEMENTOS MENOS EFICACES PARA UN APRENDIZAJE COOPERATIVO EN LA UNIVERSIDAD	283
ILUSTRACIÓN 23: IMPLICACIONES EDUCATIVAS DEL APRENDIZAJE COOPERATIVO	292





*Creo que vale la pena tratar de saber algo acerca del mundo,  
aunque al intentarlo sólo lleguemos a saber que no sabemos mucho.  
Tal estado de culta ignorancia podría sernos de ayuda  
para muchas de nuestras preocupaciones.  
Nos haría bien a todos recordar que,  
si bien diferimos bastante en las diversas pequeñeces que conocemos,  
en nuestra infinita ignorancia somos todos iguales  
Karl Popper, 1972, traducción de Néstor Miguez 1991.*



## **Abstract**

**Justification of the study.** Is not recent to state that the individual learning is the most frequent one in the university, the most criticised and the most widespread in the university field (Nelson, 1986) but neither it is less true, that the cooperative learning is not a recent methodology. The historical perspective phases it, as the first step of the inside the classroom cooperation, in the French humanistic movement of the XVIII century through the postulates of mutual education of Pestalozzi and the educational proposals of Rousseau (Ovejero, 1990a) that, with the beginning of the scientific psychology of ends of the XIX, ended in the functionalism of Dewey.

The historical beginning of the study of the cooperative learning, in the 20<sup>th</sup> decade of the last century, is and indicator of the long journey that this methodology of learning had covered and at the same time, of the relevant number of investigations that have been published until our days. It began in small scale in American laboratories and was carried to the classrooms in the year 1970 (Slavin, 1991). Precisely, the years 70th, is the decade that reports the greater number of investigations focused on the university when closed colleagues to Slavin like the brothers Roger Johnson and David Johnson, joined this initiative to educate university professors (Johnson & Johnson, 1999d) but this impulse was not keep in the next decades.

Johnson, Johnson & Smith (2007) are the first ones admitting that it is at the end of the years 90th –decade that coincides with the beginning of the study of this methodology in Spain–, when the cooperative learning returns again to the American university after a significant pool of investigations focused on other educational levels but, with the problem to be a pedagogical strategy well researched although underused in the university classrooms (Jones & Jones, 2008; Paulson & Faust, 2008).

However, a big number of the systematic reviews and meta-analysis that confirm the effectiveness of this methodology in comparison with other traditional methods regarding the achievement and attitudes towards learning (Johnson & Johnson, 1989; Slavin, 1995) have been carried out in primary school, secondary school and/or high school being scarce the quantitative syntheses oriented to study the effects of the cooperative learning in comparison with the individual learning in the academic achievement of university students.

In consequence, there are many reasons that justify the realisation of the current study. The first one is the educational level in where a lower number of quantitative syntheses exist when it is being compared with other educational phases. The second one, it has passed 10 years already from the last quantitative review. The third one, the study of other explanatory variables could provide an insight to new findings on the best conditions of this methodology in the academic achievement of university students. The fourth and last one, the demand of effective active methodologies supported on the evidence based on the educational research, is a permanent challenge for the improvement of the educational quality.

Therefore, the historical future in North America, combined with the demands of the European space of higher education and the increase educational innovation experiences in the Spaniard university, represents an international and national panorama dedicated to the development of educational models with active methodologies of education.

These demands in matter of innovation and educational research are the ones that have the purpose of educating, on one hand, a student capable of responding to the challenges that the society demands as an active citizen as well as a future professional and, on the other hand, to a university professor that improves the processes of learning through the research of new strategies and educational methodologies, as a result of a robust didactic reflection (Apodaca, 2009; Rosales, 2009).

It is the cooperative learning as learning methodology, an answer although not the only one, to a so complex reality as it is the challenge of the educational innovation, because it fulfils two fundamental conditions: it increases the academic accomplishment and besides it offers benefits in the development of personal attitudes and social skills.

**General and specific objectives of the study.** The general objective of the present study has been to prove statistically if the cooperative learning is effective as a learning methodology when it is being compared with the individual learning in the academic achievement of university students, and if it is efficient, under which conditions, being the specific objectives:

1. To integrate quantitatively the results of the cooperative learning and of the individual learning of experimental studies and quasi-experimental ones, that measures the academic achievement of university students,

2. To analyse the moderating variables of the cooperative learning that favour the academic achievement of university students,
3. To offer methodological orientations and effective elements for the put into practice of the cooperative learning in higher education.

**Method.** Defined the problem and therefore, proposed the objectives of the study, they fulfilled their own phases of all meta-analysis that are being summarised in the review of the literature, the criteria of inclusion/exclusion for the selection of the primary studies, the analysis or integration of the evidence of the primary studies, and finally, the interpretation and presentation of the cumulated evidence (Cooper & Hedges, 1994; Glass, McGaw & Smith, 1981).

A total of 11 primary sources were reviewed, conformed by 3 databases (*Web of Knowledge* –since 2014 *Web of Science*–, *Periodicals Archive Online* and *Proquest Dissertation Publishing*); 2 search engines (*Google Scholar* and *Google advanced search*) and 6 specialised journals (*European Journal of Social Psychology*, *College Teaching*, *Higher Education Research Development*, *Journal of Educational Psychology*, *Research in Higher Education Journal* and *Journal College Science*) and 7 secondary sources, namely: 3 meta-analysis, 3 primary studies and a narrative review under the descriptors of *Cooperative Learning* and *Collaborative Learning*.

The criteria of inclusion for the selection of the primary documents can be sum up in studies (a) written in English or in Spanish language, (b) published between the years 1980 until June of the year 2012 (both included), (c) with university students of any career that are studying undergraduate degrees, a diploma or a master (specialisations, master and/or doctorates), (d) where the independent variables have been the cooperative learning and the individual learning, (e) and the dependent variable, the academic achievement, which has been measured through different quantitative tools, (f) with at least an inactive control group (individual learning) and an experimental group (cooperative learning), (g) with studies that contemplate measures before and afterwards or only after the intervention and (h) those which have the sufficient statistical information for the calculation of the effect size.

For the definition and coding (manual) of the characteristics of the studies a total of 16 variables were considered for its later validation as potential explanatory factors of the results: extrinsic variables (years of publication, types of document, place of origin);

ecological variables (areas of knowledge, forming of the groups, number of members, length of the intervention, cooperative methods, structure of reward, structure of task, equality of opportunities for the group grading) and methodological variables (allocation, equivalent groups and non-equivalent groups, effects of the researcher, tools of measure and methodological quality).

Specifically, the methodological quality of the studies was assessed by a scale, designed and validated by 35 experts (19 national specialists and 16 internationals) which obtained a Cronbach coefficient of 0,941 and a moderate Kappa coefficient of 0,532 (inter-observant reliability) composed by 44 items distributed among 5 dimensions: construct validity, internal validity, external validity and statistical conclusion validity.

Regarding the measures of the results, the unit of analysis was delimited defining what is understand as a study to attend the problems of independence which faces every meta-analysis (independence between authors and research teams, independence between measures, independence between samples, independence between studies).

The effect size was calculated according to Hedges and Olkin (1985) using the standardised mean difference  $d$  between the experimental group and the control group and its interpretation followed the recommendations of Cohen (1977). The mean weighted effect size was calculated as a weighted average of the estimators of each study, using as weights  $w$  the inverse of variance for each independent effect size with a confidence interval of 95% whose results have been presented in the *forest plot*.

The fixed effects model has guided the analysis of the subgroups due to the high heterogeneity found and because in this model the only variability assumed in the individual studies, is the owed to the error of the random sampling that involves that the differences of the effect size between the individual studies are due to the fact that the investigations use samples of different subjects. Therefore, the generalisation of the results limits to the population of the studies with similar characteristics to the ones included in the meta-analysis.

The Q statistic has been used for the evaluation of the existence or not of the heterogeneity of the studies, but the  $I^2$  index has also been calculated, the one proposed by Higgins and Thompson (2002) because it estimates the degree of this heterogeneity expressed in percentage terms.

The sensitivity analysis allowed to studied the influence of each study in the result of the meta-analysis and in consequence, if the findings could see themselves affected or not by several reasons.

The bias of publication understood as the selective publication of studies based on its findings was attended through graphic methods such as the *funnel plot* analysis and analytical methods with the Egger's Test and also the fail-safe N of Orwin (1983).

**Results (sample description).** Of a total of 3744 studies registered under the descriptors *Cooperative Learning* and *Collaborative Learning* in the primary and secondary sources selected, 2536 were found, of which 2446 have been excluded and only 90 have fulfilled the inclusion criteria that report at the same time 119 outcomes in where 8114 university students have participated.

The studies between the years 90-99 prevail (33,33%), from primary sources (58,89%), research articles (75,56%) and of America specifically of the United States (67,78%), in where it has been assigned at random to the subjects and/or groups (65,55%), the initial equivalent groups and non-equivalent groups have been controlled (80%), the tools of measure applied have been standardised (51,64%), the level of methodological quality has been average (57,77%) and in science subjects (41,11%), with cooperative groups between 4 and 6 members (45,58%) but it also true that a significant number of studies does not specify if the intervention has been carried out or not by the researcher (43,33%), the cooperative method used and its classification (38,66%), the forming of the groups (46,67%), the length of the program (27,78%) and the structure of reward (81,52%) the structure of the task (81,52%) and the grading system (81,52%).

**Results (quantitative and qualitative synthesis).** The global effect size on the academic achievement in university students, has been statistically significant and positive in favour of the cooperative learning rather than the individual learning, with a moderate size in the fixed effects model ( $d=0,56$ ), a  $Q$  statistically significant with  $Q=430,44$ ,  $p<0.00001$  and a heterogeneity of 73% and, in the random effects model ( $d=0,57$ ) with a  $Q=111,2$ ;  $p<0,00001$  and heterogeneity of 0%.

Of the 119 outcomes from the 90 documents studied, in 109 of them (91,59%) the effect size has been positive with individual effects that oscillate between 0,009 and 2,50 and only in 10 of the results (8,40%) have been negative in favour of the individual learning with effects that oscillate between -0,194 and -0,027 (see illustration 8).



Of these 109 *outcomes* in where the size has been positive, the 46,78% has obtained a small size (between 0,009 and 0,47,  $n=51$ ), the 24,77% a moderate size (between 0,5 and 0,797,  $n=27$ ) and 28,44% a high size (between 0,81 and 2,5,  $n=31$ ) and at the same time, almost half of it have been statistically significant, with 53 outcomes (48,62%) in favour of the cooperative learning with effect sizes that oscillate between 0,366 until 2,5 with big effects (58,49%;  $n=31$ ) followed by average effects (35,8%;  $n=19$ ) and in last place, a lower percentage of students with a small effect (5,66%;  $n=3$ ).

The analysis by subgroups evidence positive and significant effects in favour of the cooperative learning in all the moderating variables, whose sizes oscillate between 0,18 and 1,08, with the exception of the pair bonding subgroup conformed by an only study with a negative effect ( $d= -0,04$ ) but non-significant ( $IC_{95\%}= -0,27; +0,19; p=0,743$ ) and other two sub-groups that although reported small effect sizes but neither are significant: the period 2000–2004 with a  $d=0,31$  ( $IC_{95\%}= 0; +0,63; p=0,0473$ ) and the mixed measure tools with a  $d=0,18$  ( $IC_{95\%}= -0,09; +0,45; p=0,18$ ).

The results of the sensitivity analysis have been robust. The effect size when they are being eliminated oscillates between 0,55 and 0,57, what supposes a 2% less and a 2% more with regard to the global effect size of 0,56 (fixed effects model). Neither a striking change is evident in the heterogeneity index, which oscillates between a 70% and a 73%, by what can be said that none of these affects in a remarkable way the global estimated result ( $d=0,56$ ).

The *funnel plot* analysis revealed that the presence of a publication bias is refused, data that was equally confirmed by the Egger's Test,  $p=0,727$  and the fail-safe N of Orwin (1983) with a high number of non-recovered studies 214,2.

Regarding the quality level of the studies the effect does not vary. In studies with a low quality the effect size is moderated with a  $d=0,74$  and is keep between an average and high level ( $d=0,52$  and 0,58, respectively). If the global effect size in studies without quality has been moderated with 0,56 (fixed effects model) this size is equally conserved in studies with low, average and high quality although the low level has obtained the higher effect between the three of them ( $d=0,74$ ).

**Discussion and conclusions.** The results obtained offer guidance to the university professors for the put into practice of the cooperative learning under effective conditions in the academic achievement of the students.

The most efficient mechanisms, meaning, those that have obtained **big effects** in **Statistics Subjects, Chemistry, Nursing and Physics**, with the cooperative methods ***Jigsaw I, Jigsaw II and Group Investigation*** that require of a specialisation of the task to attain the aims proposed and, regarding publications, the most recent ones have obtained bigger effects specifically between the years **2010 and 2012**.

To these optimum conditions are added those moderate effects that prevail in each one of the variables studied because it deals with factors that equally concern to the university professor for the decision-making in his classroom. These moderate effects are evident in fields of knowledge like **Sciences** in subjects of Statistics, Physics, Mathematics, Chemistry and Computing also in **Health Sciences** specifically in Nursing and in **Education**, when training professors in Physical Education, Mathematical Education, Musical Education and Education for Reading and Writing; with methods focused on the use of the ***Computer Support***, in **low length programs** – between 1 and 4 weeks–, with **odd heterogeneous groups conformed between 4 and 6 members** and preferably with those **cooperative methods that require of a specialisation of the task** (*Jigsaw I, Jigsaw II and Group Investigation*), in where **rewards or group gratifications are not offer** (*Jigsaw I and Group Investigation*) and when **not all the group members have the same opportunities to contribute with their performance to the grading of the group** what translates in different qualifications for each member (*Learning Together, Jigsaw I and Group Investigation*).

When the university professor decides to assume methodological conditions, one has to be aware that moderate effects that equally predominate are the product of the **random allocation of the subjects and groups**; if it is **the same researcher** the one who carries out the intervention; if it guarantees **the equivalence of the groups** before the launch of the program and if the **tools are standardised** and that, the **technical reports** (extrinsic variable) have also stood out in the effect size although they have not differentiated significantly of other types of documents such as the dissertations and articles published in scientific journals.

Regarding the **methodological quality** and its impact, if the studies are characterised by a low level of quality although they equally keep a moderate effect in the academic achievement tends to be more elevated in comparison with studies with an average and high quality, the reason why, an analysis of those conditions that are or not affected by the quality of the studies is another factor to consider.

For example, if they give classes in **Education and Social Sciences, Commercial Education and Law** or use the *Computer Support* and *Dyadic Methods* they keep the effects in the achievement of the students without being the quality a factor that impact in the academic. The same occurs with documents published between the years **1980-1999** and between the years **2000-2012**.

Instead, **there are other conditions that are being affected by the quality of the studies** like the *Student Team Learning* method, the **combined methods** and the fields of **Science and Humanities and Arts** (illustration 18).

However, if we precise by **areas of knowledge, the most efficient conditions when the cooperative learning is applied in Sciences** –Statistics, Physics, Mathematics, Chemistry and Computer– are those that use (a) the combined methods for the physics education and the technology (computer), (b) the *computer support* to teach subjects of technology (computer), Chemistry, Statistics and Mathematics and (c) the methods that gathered under the classification of *Cooperative Investigation* that includes a *Group Investigation* and *Jigsaw I* for the education of the Chemistry.

Continuing with the analysis by fields of knowledge but now **in Social Sciences**–Administration, Economy and Psychology– the most efficient cooperative methods are those that gather under the denomination of *Dyadic Methods* for the subject of Psychology, specifically *Reciprocal Peer Tutoring* and *Think–Pair–Share*–.

Finally, **in Humanities and Arts** –Linguistic and Music– the most efficient cooperative methods are those that are based on *Computer Support* in face-to-face or virtual modality for the subjects related with the Linguistics as the method of *Cooperative Computer–Assisted Language Learning* (CALL) used in distance education for the learning of a foreign language or the face-to-face modality when it uses *E–Learning with Cooperative Learning* (ELCL) for the oral proficiency of a second tongue.

From the findings found, the recommendation is (a) to expand the sample with the incorporation of primary studies subsequent to the year 2012, the review of other databases, specially of specialised journals in Europe and in less explored subjects such as Agriculture, Economy, Engineering, Physics and Nursing; (b) to work with the term of *Collaborative Learning* in the *Web of Knowledge* database; (c) to include other specific descriptors with specific cooperative methods like *Scripted Cooperation*, *Jigsaw*, *Reciprocal Peer Tutoring*, *Student Teams–Achievement Divisions*, *Teams–*

*Games–Tournament*, among others; (d) to study in depth moderating variables that affect the operation of the cooperative groups, that although they have acquired a moderate size not negligible, non of them excel with a big effect; (e) to continue exploring in the information and communication technologies in cooperative/collaborative situations as an scenario of interaction; (f) to incorporate in the future quantitative syntheses the academic course in which the cooperative learning has been worked; (g) to continue with the incipient bibliometric analysis that has begun; (h) to study the culture of origin of the university students like another moderating variable; (i) to improve the scale for the assessment of the methodological quality of the primary studies; (j) to study the cooperative base groups and its incidence in the development of social skills; (k) the incursion in systematic syntheses of qualitative research; (l) to develop a cumulative meta-analysis to offer new findings to the current meta-analysis and finally; (m) a meta-regression with the purpose to offer a more complex model in where the most relevant characteristics of the studies could explain the percentage of variance of the effect sizes.

In summary, the findings found are a small contribution to improve the practices and university quality because when it comes to educational innovation, the education methodologies are the fundamental axis for the change, a change that should be supported based on evidence.

It is the Evidence Based on Educational Research the integration of the empirical demonstration and the continuous training of the professor's staff what enables the decision-making for the teaching of the education. In the empirical evidence falls on one hand, the weight of presenting the conditions for the implementation of effective education methods and on the other hand, the wisdom and professional ability of the professor to apply these methodologies with its group of students (Moran, 2004).



## **Introducción**

Preguntar. Preguntar una y otra vez es una forma que desde el interés por conocer permite introducir al investigador en un objeto de estudio. Producir una batería de preguntas es, tal vez, la mejor forma de dar un primer paso en materia de investigación y, en especial, cuando existe evidencia acumulada durante años en cualquier ámbito del saber. Precisamente, preguntar y repreguntar, a partir del cúmulo de esas investigaciones precedentes, es considerar que se necesita integrar de manera sistemática y bajos criterios específicos, el crecimiento exponencial de los hallazgos científicos para una toma de decisiones fundamentada en la investigación basada en la evidencia.

Preguntar nuevamente sobre la base de la colección de estudios individuales que se han reunido en torno al aprendizaje cooperativo, es la manera de adentrarse en la lectura de estas páginas, que buscan acercarse a una metodología de enseñanza–aprendizaje que no es reciente y que aunque guarda en su haber mas de 40 años de producción científica, el contexto universitario es el ámbito que menos estudios reporta. ¿Es el Aprendizaje Cooperativo eficaz como metodología cuando se compara con el aprendizaje individual? De ser así, ¿bajo qué condiciones el Aprendizaje Cooperativo mejora el rendimiento de los estudiantes universitarios? Y, por tanto, ¿puede trabajarse como metodología activa tal y como exige el Espacio Europeo de Educación Superior para la mejora de la práctica educativa?

Así, el aprendizaje cooperativo se convierte en el objeto de estudio de esta tesis, en el entramado necesario. Su historia, las razones de irse constituyendo en metodología de apoyo a la educación, los elementos que lo conforman, los factores y condiciones que lo dibujan, la versatilidad de sus técnicas, sus beneficios y limitaciones, es ineludible fuente de revisión dentro de la Universidad.

Necesario será entonces remontarse a los orígenes del aprendizaje cooperativo, ir hasta el siglo XVIII como espacio temporal en el que surgen los primeros postulados de Pestalozzi y Rousseau. Es cuando se empieza a trabajar el concepto de enseñanza mutua, como proceso en donde se busca que sea el alumno el que defina junto a un compañero sus relaciones con el aprendizaje bajo una meta en común, son el hito para que pensadores y educadores inicien una labor de creación de métodos de enseñanza cooperativas.

Luego, ya en el siglo XX serán los aportes de la Teoría del Aprendizaje por Comportamiento que se centra en el refuerzo y retribución del grupo, de Piaget y su conflicto cognitivo, de los neopiagetianos y el conflicto sociocognitivo, de Vygotsky y el interaccionismo social, de la Teoría de la Interdependencia Social en sus inicios con Kufka y que es retomada por los hermanos Johnson para seguir profundizando y continuar con la Teoría de la Controversia Constructiva, los que van conformando los cimientos psicopedagógicos del aprendizaje cooperativo.

Este paseo teórico no se detiene, sino que se emprende desde nuevas miradas dando pie a una contemporaneidad que se muestra en la curiosidad investigadora internacional con ejemplos como los de Slavin, los hermanos Johnson y Kagan, quienes desde la década de los setenta estudian en Estados Unidos a esta metodología de enseñanza. O en España, con estudiosos del aprendizaje cooperativo como García Hoz, Medina Rivilla, Ovejero, Coll y Corominas, entre otros muchos que también tienen un recorrido que parte en las mismas décadas de los setenta y ochenta del siglo pasado, pero que es realmente en los noventa cuando se comienza a estudiar la cooperación y su implementación en las aulas españolas.

Es esa acumulación de investigaciones empíricas una consecuencia de la utilización de términos similares o próximos al aprendizaje cooperativo como trabajo en grupo, trabajo en equipo o grupos de trabajo, entre otros, en donde los límites entre cada una de esas definiciones en ocasiones es confuso y, aun mas cuando se ubican a todos debajo del gran paraguas de la cooperación.

Sin embargo, son dos las acepciones que mayor debate presentan en la comunidad científica, porque aunque ambas parten del constructivismo, se diferencian principalmente en su grado de estructuración, en el rol del alumno y en especial de la función del profesor. Es el aprendizaje cooperativo y el aprendizaje colaborativo un ejemplo de diferencias sutiles y no tan sutiles que han colmado las páginas de revistas científicas, siendo el primero el que cuenta con el mayor volumen de investigaciones en todas las etapas educativas, no solo en cantidad, sino también en la especificidad de las diversas técnicas cooperativas, aunque es cierto que el aprendizaje colaborativo es el que se ha diseñado para el ámbito universitario.

Visto lo anterior, la frontera entre acepciones análogas es una discusión que se entrelaza con la comprensión de los distintos tipos de aprendizaje que ocurren en el aula y que

tampoco pueden escabullirse del contexto histórico, social y cultural en el que se produce la enseñanza. En consecuencia, el aprendizaje cooperativo frente al aprendizaje individual y al aprendizaje competitivo son polémicos en si mismos, no porque sus objetivos sean difusos sino por sus beneficios en el aprendizaje –unos a favor y otros en contra– y en especial, por las implicaciones educativas que cada uno de estos tres tipos de aprendizaje exige al profesorado.

El aprendizaje cooperativo propone que sean los propios alumnos con sus diferencias y desde sus experiencias, los que emprendan la tarea de enseñar y de aprender de sus compañeros. Requiere entonces de parte del docente una preparación de las clases de modo diverso que favorezca la interacción entre estudiantes y que promueva el conflicto. Una estructura de los contenidos perfectamente atados, diseñados para que el alumno pueda encontrar y plantearse que los contenidos de aprendizaje son, por decirlo de alguna forma, un hallazgo, una forma de hacer y de entender.

En cambio, el aprendizaje individual hace que el estudiante sea el único artífice de su proceso de aprendizaje. El desempeño de cada estudiante no se relaciona y no depende de los demás. Los alumnos buscan sus beneficios personales y por consiguiente, no se establece una relación de interdependencia con sus compañeros para lograr una meta común

En contraste, el aprendizaje competitivo –el menos estudiado de los tres– busca que los alumnos compitan entre sí, procurando superarse a sí mismos, con el concurso de una rivalidad externa generada por los demás.

Quienes abogan por el aprendizaje cooperativo como una metodología a utilizar en el aula universitaria –sin que ello implique el uso exclusivo de este tipo de aprendizaje–, parten de la necesaria existencia de un grupo que busca en el compartir, en la discusión que emerge del conflicto, alcanzar la meta del conocimiento en donde las habilidades sociales son el mecanismo para lograr el éxito académico. Son elementos fundamentales de un grupo cooperativo eficaz, la interdependencia positiva, la interacción cara a cara, las habilidades interpersonales y en grupo y el procesamiento y la evaluación grupal. Todos estos principios reflejan una enorme responsabilidad de cada uno de los participantes del grupo, de la buena preparación de las clases por parte del docente y es, en esa estrecha relación en conjunto, cuando todas las condiciones y variantes están suficientemente atadas para y por la labor de la enseñanza a favor del aprendizaje.



Cuando se garantizan todos estos elementos esenciales, porque la falta de uno de ellos convertiría la situación cooperativa en una práctica común en las aulas bajo el nombre de “trabajo en grupo” –que dista significativamente de las características propias de este tipo de aprendizaje–, la eficacia del aprendizaje cooperativo es superior cuando se compara con el aprendizaje individual y competitivo, y mas aun, esta eficacia se incrementa cuando se asumen variables explicativas relacionadas con los procesos cognitivos, sociales e instructivos que ocurren en la interacción.

Evidentemente, cada aula, cada asignatura, la composición y el tamaño de los grupos, las competencias intrapersonales e interpersonales que se desarrollan en la interacción, el sistema de recompensas y los propios intereses docentes, han generado diferentes variantes para la construcción de abundantes métodos cooperativos. Variantes que han dado pie a formas de organización del aula muy versátiles que han redundado en un soporte de posibilidades diversas sobre un mismo objeto: que el estudiante interactúe con otro para mejorar su formación.

Es sabido que la eficacia del aprendizaje cooperativo se ha medido en tres grandes dimensiones, a saber: el rendimiento, las relaciones interpersonales y el ajuste psicosocial que se establece durante la interacción. Así, el rendimiento académico es la dimensión que mayor volumen de trabajos publicados alberga debido al interés que siempre ha suscitado en Psicología y en Educación, pero a la vez, se sigue manteniendo la tendencia de ser el contexto universitario el que cuenta con el menor número de estudios cuando se compara con el individual y, todavía en un menor grado con otras etapas educativas.

Así, centrándonos solo en los meta-análisis que se han llevado a cabo en la universidad, en donde se confronta el rendimiento académico en distintas situaciones de aprendizaje, como es el aprendizaje cooperativo y el individual, la eficacia de lo cooperativo es superior, planteando que si bien es cierta esta afirmación, la investigación que se ha desarrollado en el ámbito universitario es mucho más reducida que los estudios en otros niveles de enseñanza y, por consiguiente, no es la etapa educativa que cuenta con estudios meta-analíticos recientes.

Y ahí está el propósito, el objeto de estudio de esta tesis. Se trata de incorporar estudios previos utilizados en otros meta-análisis y de sumar investigaciones actuales que cumplan con los criterios de inclusión para comprobar estadísticamente si el aprendizaje

cooperativo es eficaz como metodología de aprendizaje cuando se compara con el individual y, de ser eficiente, bajo qué condiciones mejora el rendimiento académico de estudiantes universitarios. De esta manera, se valorarán los elementos mas eficientes que hacen que la formación adopte la característica de ser cooperativa, colectiva y conviertan al aula en una comunidad de aprendizaje.

Las metodologías de enseñanza–aprendizaje centradas en el alumno –nombradas por algunos autores como activas– en donde el aprendizaje cooperativo es una de ellas, se encuentran en plena ebullición de difusión, uso y análisis y hacen que las aulas de clase estén convertidas en laboratorios de experimentación en materia de aprendizaje. Y tiene sentido esto.

Los llamados ‘grados universitarios’ –adecuados a las directrices del Espacio Europeo de Educación Superior– exigen al profesorado una renovación de las metodologías de enseñanza y una formación continua en donde la innovación educativa esté al servicio de la calidad. Las universidades ya han generado egresados y estos están empezando a ocupar puestos de trabajo y es, precisamente ahora, el momento oportuno para afinar y adecuar las estructuras de enseñanza, porque el trabajo en grupo es una competencia transversal en todas las titulaciones y una competencia indispensable en el mundo laboral.

Estudiar al aprendizaje cooperativo es tratar de dar respuesta a la mejora del rendimiento de los estudiantes universitarios bajo las condiciones mas eficaces amparada en una Evidencia Basada en la Investigación Educativa en donde se trabajan habilidades sociales tan necesarias en una sociedad democrática.

En Educación, la mayor trayectoria de estudios meta–analíticos proviene fundamentalmente de los Estados Unidos de América en donde existe una larga tradición en indicadores de calidad para la valoración de los estudios. Por su parte, en Europa, 23 de sus países se han interesado en la Evidencia Basada en la Investigación Educativa para la formulación de sus políticas educativas. Y, en España mientras esta investigación basada en la evidencia reclama por su consolidación, hay quienes comienzan a defender una investigación sustentada en evidencias científicas para la enseñanza universitaria.

Es el meta–análisis el método adecuado cuando se necesita sintetizar y valorar un número importante de estudios en un área de conocimiento, como lo es el aprendizaje

cooperativo, con el fin último de integrar las conclusiones para una toma de decisiones educativas fundamentadas en la evidencia.

De esa manera, la tesis está estructurada en dos grandes apartados que se entretajan. La *parte teórica*, la urdimbre relativa a los fundamentos del aprendizaje cooperativo. La *parte empírica*, la trama que se va hilando en los cimientos teóricos y que aglutina al método, los resultados y la discusión para cerrar con conclusiones e implicaciones educativas a la que se suman sus respectivas limitaciones y prospectiva.

En su primera parte, *la teórica*, el recorrido se inicia en los fundamentos que dan vida al aprendizaje cooperativo y cómo se ha ido nutriendo esta metodología de enseñanza–aprendizaje desde la psicopedagogía. Luego, se definen los postulados teóricos que, por contraste con el aprendizaje cooperativo, tienen el aprendizaje individual y competitivo, para definirlo a su vez desde la oposición con el aprendizaje colaborativo.

Un especial hincapié se hace sobre las variables moderadoras del aprendizaje cooperativo, como son las cognitivas, sociales e instructivas, que a su vez condicionan y dan forma a las diferentes técnicas de enseñanza, los métodos que se han ido gestando y desarrollando para la buena práctica de esta metodología de enseñanza–aprendizaje. El trayecto en ese sentido da cuenta de los métodos del *Student Team Learning* y sus diversas variantes, del *Cooperative Investigation* y sus derivados, del *Jigsaw* y sus múltiples posibilidades, de los *Dyadic Methods* basados en la interacción entre parejas, del método *Learning Together*, los *Informal Methods*, y del aprendizaje cooperativo asistido por ordenadores.

En la misma revisión teórica, las *evidencias empíricas* son el alma de esta parte. Las investigaciones desarrolladas sobre la base del aprendizaje cooperativo en universidades internacionales y españolas, muestran el balance de unos estudios que plantean la eficacia del método en tres grandes resultados: el esfuerzo ejercido para alcanzar el logro, la calidad de las relaciones, el ajuste psicosocial y las competencias sociales. Estos buenos resultados son una consecuencia de las relaciones recíprocas entre dos factores que definen al aprendizaje cooperativo: la interacción positiva y la promoción de interacción entre los participantes.

Las *perspectivas internacional y española en investigaciones sobre aprendizaje cooperativo en la universidad* son un menú de las opciones sobre las que los investigadores en los últimos años se han ocupado. Pequeños grupos, juegos de rol,

beneficios para el aprendizaje y retención de lo aprendido, la promoción del pensamiento crítico, la integración social, la innovación educativa, entre otros, son parte de esas evidencias que permiten reconocer al método como valioso dentro de la educación.

Cerrando el apartado teórico, la *eficacia del aprendizaje cooperativo en el rendimiento: meta-análisis*, es un capítulo que plantea la necesidad de la integración de la investigación a través de un análisis estadístico para la toma de decisiones en Educación y por ello, hace una compilación de los estudios meta-analíticos que se han realizado en la universidad.

Posteriormente, se menciona el debate actual en torno a la investigación educativa, a la llamada *Evidencia Basada en Investigación Educativa*, que ubica al método más cerca del objeto de estudio de esta tesis y que muestra las fortalezas, pero también las dificultades a las que se enfrentan los investigadores cuando tienen que establecer los criterios para la selección y posterior análisis de las “mejores evidencias en Educación”. El reto, plantea la necesidad de continuar trabajando en el meta-análisis, llamado también “el análisis del análisis” y creado por Glass en los años 70 para una investigación al servicio de la práctica educativa.

Los estudios *meta-analíticos en los que se incluye el aprendizaje cooperativo en la Universidad* ubican al lector en el devenir y avances de este tipo de investigaciones estadísticas, aunque el nivel educativo universitario de estos estudios no es el protagonista. Importantes hallazgos en esa materia son descritos en esta parte. La relación entre agrupamiento o no, las facilidades para el rendimiento estudiantil que proporciona el aprendizaje cooperativo máxime si se contrasta con el rendimiento individual, las comparaciones entre grupos heterogéneos o no entre otros, son los resultados de esta revisión teórica.

Más adelante, en el apartado *meta-análisis centrados solo en la universidad*, en donde la universidad si es la protagonista, las investigaciones en aprendizaje cooperativo muestran la enorme variabilidad de disciplinas que se dictan en la universidad y cómo los métodos de enseñanza van conformando un andamio en el que, la aplicación de metodologías cooperativas mejora el desempeño en contraste con metodologías centradas en el profesor, como es el caso del aprendizaje individual. A pesar de ser la clase magistral el método de enseñanza preferido y por consiguiente, el aprendizaje

individual el mas utilizado en la universidad, los beneficios de la cooperación en el rendimiento académico son potentes.

Ya en la *parte empírica*, además de justificar el estudio y darle el cauce a través de la definición de los objetivos, se adelanta la necesidad de la comprensión y uso del meta-análisis como realidad que exige al docente proponer desde su labor de enseñanza, rendimientos superiores a sus estudiantes y prepararles desde las competencias del trabajo cooperativo.

En el apartado del *método*, teniendo muy claro y sin perder de vista el objetivo del estudio, se aplican las fases propias de un meta-análisis. Luego de la revisión de la literatura y las estrategias empleadas para su selección, así como de los criterios de inclusión/exclusión de estudios primarios a partir de las fuentes primarias y secundarias consultadas, se procede a mostrar la codificación y el manual que se ha creado *ad-hoc* para la valoración de la calidad metodológica de los estudios primarios en sus tres fases: construcción y diseño, juicio de expertos y fiabilidad inter-observadores.

Por último, ya cerrando el método, se muestran las decisiones metodológicas asumidas con relación a la *medida de los resultados*. Se menciona el *software estadístico* usado, la *unidad de análisis estadístico* para atender el problema de independencia, la *magnitud del efecto global*, la *estimación global del efecto* y la justificación de la elección del modelo de efectos fijos, la *heterogeneidad de los estudios*, las *variables moderadoras* como posibles fuentes de explicación, el *análisis de sensibilidad* y, por supuesto, el *sesgo de publicación*.

Es el análisis de los resultados de los estudios primarios, así como la interpretación de la evidencia acumulada y el propio interés en comprobar si el aprendizaje cooperativo es una metodología eficaz frente al aprendizaje individual entre estudiantes universitarios, las razones principales en las que se sostiene *los resultados y la discusión*.

Es la *descripción de los estudios*, la *síntesis cuantitativa* que ofrece la *magnitud del efecto global del rendimiento*, el *análisis de sensibilidad*, la *heterogeneidad* de los estudios, el *análisis por subgrupos* y por *estratos*, el *sesgo de publicación* y la *síntesis cualitativa* y esa necesaria discriminación relativa a la *calidad de los estudios*, los que dan pie finalmente a la comprobación de los objetivos de partida a través de una discusión que decanta en los *elementos eficaces para un aprendizaje cooperativo en la universidad*.

Las *conclusiones* como proposición final junto con las implicaciones educativas invitan al lector a reflexionar sobre las relaciones que se producen al aplicar la metodología del aprendizaje cooperativo en el ámbito universitario como innovación educativa que involucra a estudiantes, profesores y facultades y que permea hacia el ámbito empresarial.

Conocer *las limitaciones* y *las fortalezas* a partir de los hallazgos encontrados y con una visión a corto y a mediano plazo de hacia dónde se dirigen las investigaciones en aprendizaje cooperativo, plantea nuevos caminos a seguir en donde **preguntar una y otra vez permitirá introducir al investigador en un nuevo objeto de estudio por explorar.**



## **PARTE TEÓRICA**





## **Capítulo I. APRENDIZAJE COOPERATIVO: FUNDAMENTOS TEÓRICOS**

### ***1.1. Teorías psicopedagógicas del aprendizaje cooperativo***

Estudiar las Teorías del Aprendizaje Conductual, la Teoría del Desarrollo Cognitivo, el papel que juega el conflicto cognitivo y el conflicto socio–cognitivo en el aprendizaje, la Teoría Sociocultural, la Teoría de la Interdependencia Social y la Teoría de la Controversia Constructiva, es reflexionar sobre el devenir de la enseñanza y ver cómo estas teorías psicopedagógicas inciden en la manera de comprender las relaciones entre pares, el rol del profesor y del alumno, el proceso de enseñanza–aprendizaje y por consiguiente, ahondar en sus implicaciones en la aplicación del aprendizaje cooperativo como metodología de enseñanza.

#### **1.1.1. Teorías del Aprendizaje Conductual**

Hoy en día el conductismo sigue siendo una de las teorías mas determinantes en la psicología educativa, afirmación que no invalida el hecho de que sus bases explicativas sobre cómo aprende la persona estén superadas por otras teorías aunque se siguen manteniendo algunos de sus principios sobre el aprendizaje, lo cual explicaría el por qué el aprendizaje cooperativo como metodología de enseñanza funciona y el por qué continúan destacando sus implicaciones en el contexto educativo (Moruno, Sánchez y Zariquiey, 2012).

La perspectiva del aprendizaje conductual “sostiene que los estudiantes trabajarán arduamente en tareas en las cuales ellos tengan asegurada una recompensa de algún tipo y fracasarán al trabajar en tareas en las cuales no haya una recompensa o bien la tarea conduzca a algún tipo de castigo” (Bandura, 1977 y Skinner, 1968 citados por Johnson, Johnson & Smith, 1997, p. 10). Entonces ¿cuáles son los aportes de Skinner y Bandura para el aprendizaje cooperativo? El primero se centró en las contingencias grupales y el segundo, en el papel de la imitación y la influencia social en el aprendizaje.

El refuerzo es un elemento clave en la Teoría Conductual de Skinner, la cual se sustenta en la idea de que el aprendizaje es una función que cambia en el comportamiento. Cuando un patrón de estímulo–respuesta es reforzado, el individuo está condicionado

para responder. Por tanto, el refuerzo y las recompensas son fundamentales para incrementar el aprendizaje (Agarwal & Nagar, 2011).

En la clase tradicional, la recompensa positiva del comportamiento usualmente proviene solo del profesor. Esta situación de aprendizaje promueve una interdependencia social negativa entre los estudiantes quienes compiten entre sí para lograr la gratificación del docente. En contraste, en el aprendizaje cooperativo se promueve la interdependencia social positiva entre los compañeros, quienes son una fuente de reforzamiento para el aprendizaje. Trabajando cooperativamente ellos aprenden el material al cuestionar, explicar, recordar y asistir a los otros durante el proceso ofreciendo retroalimentación y recibiendo recompensas. Son precisamente esas recompensas las que incitan a los estudiantes en el uso de habilidades de razonamiento, a trabajar duro por el éxito y a ayudarse entre ellos (Agarwal & Nagar, 2011).

La utilización o no de recompensas extrínsecas en situaciones cooperativas es un debate que aun persiste en el escenario académico. Slavin y Kohn son un ejemplo de posiciones contrarias. El primero, es de la postura que para incrementar el desempeño en los grupos cooperativos es necesario que la recompensa grupal se sustente sobre el rendimiento individual de cada miembro del grupo. El segundo, considera que nunca debería usarse recompensas externas porque socava la motivación intrínseca para el aprendizaje “the ‘social rewards’ of working cooperatively probably enhance intrinsic motivation, and are among the great advantages of using cooperative learning strategies” (Graves, 1991, p. 77).

Por su parte, Bandura, desde la Teoría Social del Aprendizaje concibe el aprendizaje en grupo desde el papel de la imitación y la influencia social. Dos citas textuales reflejan su perspectiva. La primera, refiere a la influencia social “Human expectations, beliefs, emotional bents and cognitive competencies are developed and modified by social influences that convey information and activate emotional reactions through modeling, instruction and social persuasion” (Bandura, 1989, p. 3). La segunda, hace mención a la organización y estructura de la clase porque “Classroom structures affect the development of intellectual self-efficacy, in large part, by the relative emphasis they place on social comparison versus self-comparison appraisal” (Bandura, 1994, s.p.). En la estructura de clase cooperativa los estudiantes trabajan juntos y se ayudan entre si para promover autoevaluaciones mas positivas y alcanzar altos niveles de rendimiento lo que no ocurre en situaciones competitivas o individuales.

Es en el aprendizaje entre iguales donde ocurren verdaderos aprendizajes significativos. Es en la interacción entre un modelo y un aprendiz con rasgos similares cuando se potencia el aprendizaje vicario a través de procesos básicos como la atención, la retención, la producción y la motivación. Por medio de la observación, los estudiantes en situaciones cooperativas, aprenden nuevas conductas que no estaban inicialmente en el repertorio de conductas individuales y sus consecuencias (efecto de modelado), fortalecen o debilitan conductas aprendidas en función de las consecuencias positivas o negativas del modelo (efectos inhibitorios o desinhibitorios) pero también facilita la puesta en marcha de patrones previamente aprendidos (efecto de facilitación) (Moruno, Sánchez y Zariquiey, 2012).

### **1.1.2. Piaget y el conflicto cognitivo**

Una de las grandes teorías que explican el desarrollo cognitivo, es la Teoría del Desarrollo de la Inteligencia de Jean Piaget (1950, 1986), para quien el progreso intelectual depende en gran medida de la relación desequilibrio–equilibrio más que de la interacción social porque es este uno de los factores del desarrollo (maduración, experiencia y ambiente social), subordinado al motor interno y autónomo del mismo (Martí, 1991).

Piaget (1928 citado por DeVries, 1997) insistió en que la vida social es indispensable para el desarrollo de la lógica. El progreso en lo social y en la lógica van complementándose y constituyen aspectos indisociables de una única realidad que es a la vez social e individual. No es posible hablar de causa y efecto sino que debe verse como aspectos que se desarrollan en espiral. Es el mismo autor quien se pregunta si los niveles de la estructura lógica o prelógica son los que determinan el correspondiente modo de colaboración social o si son las estructuras de las interacciones sociales las que delimitan la naturaleza de las operaciones intelectuales.

La cooperación se define como “el esfuerzo para obtener metas comunes al tiempo que se coordinan los sentimientos y la perspectiva propia con una conciencia de los sentimientos y las perspectivas de los otros” (Piaget, 1950 citado por Johnson, Johnson y Smith, 1997, p. 9). La operación y la cooperación son conceptos inseparables de una misma realidad aunque Piaget en sus trabajos enfatizó cómo las operaciones permiten la cooperación pero no explicó cómo la cooperación hace posible las operaciones (Martí, 1991) porque

Todo el pensamiento lógico es socializado puesto que implica la posibilidad de comunicación entre individuos. Pero tal cambio interpersonal se lleva a cabo a través de correspondencias, reuniones, intersecciones y reciprocidades, es decir, a través de operaciones (...) Las acciones, ya sean individuales o interpersonales, están en esencia coordinadas y organizadas por estructuras operacionales construidas espontáneamente en el curso del desarrollo intelectual (Piaget, 1924/1993, pp. 214–215).

Es en las relaciones entre iguales en donde ocurren verdaderas oportunidades para practicar la reciprocidad –elemento básico de la justicia–, la cooperación y para concebir la justicia o la injusticia de ciertas experiencias, principios fundamentales de una moral autónoma. La cooperación entre los individuos es precisamente la que induce al respeto mutuo y la igualdad (Piaget, 1983).

Son los propios compañeros quienes facilitan ocasiones, desde esa condición de otro, para que el desarrollo cognitivo surja a partir del conflicto cognitivo porque son las diferencias de un conocimiento previo respecto al mismo lo que podría generar ese conflicto. La interacción social es un catalizador para que los estudiantes reevalúen sus creencias acerca del mundo. Si la relación entre los pares es socialmente simétrica será mas beneficiosa porque una verdadera negociación del razonamiento no ocurre entre aquellos que son expertos o que tienen distintas posiciones de autoridad dentro del grupo (Hogan & Tudge, 1999).

La confrontación de puntos de vista divergentes es lo que lleva al individuo, por una parte, a un conflicto social, que le ayuda a la mejora de la comunicación y a tomar conciencia del otro por las diversas posibilidades de puntos de vista y, por la otra, a un conflicto cognitivo, porque conduce al sujeto a plantearse que sus puntos de vista pueden ser examinados sobre la base de los conocimientos que se adquieren del otro (Torrego y Negro, 2012).

Aunque Piaget se centró en el estudio de las acciones internas –mecanismos endógenos– que permiten el desarrollo de las operaciones mentales y no en las acciones externas –entorno social– (Garton, 1994), no es menos cierto que las implicaciones educativas de su teoría –constructivismo– se traducen en que el aprendizaje entre pares constituye un núcleo relevante en la dinámica de clase porque: (a) favorece el desarrollo de estructuras intelectuales superiores como el razonamiento, la planificación, la creatividad, la resolución de conflictos y la toma de decisiones; y (b) repercute en avances cognitivos significativos al movilizar las estructuras cognitivas existentes y

obligar a reestructurarlas dando lugar al progreso intelectual (Laboratorio de innovación educativa, 2009).

### **1.1.3. Los neopietistas y el conflicto sociocognitivo**

Pero ¿por qué los neopietistas, provenientes de la escuela de Ginebra y seguidores de Piaget no asumieron el conflicto solo como cognitivo?

Willen Doise, Anne-Nelly Perret-Clermont y Gabriel Mugny, modificaron el concepto piagetiano de conflicto cognitivo, introduciendo la dimensión social, elaborando así el concepto de “conflicto socio-cognitivo” para explicar cómo cambia la comprensión del individuo al interactuar con otro u otros que tienen una concepción distinta de los hechos. Apuntan “puesto que la causalidad de lo social ha sido ignorada excesivamente, la psicología social genética, concretamente a través de esta teoría del conflicto sociocognitivo, se ha encargado de devolver a la posición que en realidad ocupa el desarrollo sociocognitivo” (Carugati y Mugny, 1988, p. 94).

Cuando dos visiones distintas del mundo entran en contacto para resolver un conflicto, este proceso hace que se produzca una “reestructuración cognitiva” desencadenando algún tipo de aprendizaje y de comprensión: “Así pues, la interacción social y conflictiva puede ser considerada como estructuradora y generadora de nuevos conocimientos” (Carugati y Mugny, 1988, p. 81).

Existen dos razones por las cuales la heterogeneidad de las respuestas genera progresos cognoscitivos. La primera, es que el individuo toma conciencia de que existe la posibilidad de otras respuestas diferentes a las de él, lo cual crea un conflicto sociocognitivo porque promueve el conflicto individual y constituye una fuente de descentralización

El desacuerdo entre los partenaires no es sólo un simple desacuerdo respecto a las respuestas cognitivas divergentes; proviene también de las relaciones entre los individuos que se encuentran al mismo tiempo con que se les ha planteado un problema social (Carugati y Mugny, 1988, p. 89).

La segunda razón, es que cuando algún miembro del grupo plantea sus ideas, ayuda a los otros a elaborar una respuesta nueva. Los beneficios del conflicto sociocognoscitivo no residen en la actividad individual cognitiva, sino que la actividad misma involucra a la persona en una relación social específica con el otro (Carugati y Mugny, 1988).

En el conflicto sociocognoscitivo no existe primacía de lo social sobre lo individual. Para los neopiagetianos, lo social y lo individual se mueven y se transforman a lo largo del desarrollo de forma recíproca. El término “sociocognoscitivo” posee dos elementos: lo social, porque surge de la información, ideas u opiniones de la interacción entre dos o más personas, y de lo cognoscitivo, porque se ubica en un plano intraindividual (Enesco y Del Olmo, 1992).

Desde la mirada del aprendizaje cooperativo, Piaget y la escuela de Ginebra –con sus respectivos énfasis en lo cognitivo y en lo social– comparten que la “confrontación de puntos de vista moderadamente divergentes” (Ovejero, 1990a, p. 67) produce en el alumno un conflicto socio–cognitivo que es el movilizador del aprendizaje, ergo, del progreso intelectual (Ovejero, 1990a).

Sin embargo, es sabido que los neopiagetianos no se dedicaron a estudiar en profundidad la conversación real implicada en estos conflictos de ideas ni tampoco el diálogo ni la discusión emergente de los grupos porque el interés se centró en determinar si esa interacción social mejoraba a posteriori las realizaciones individuales y no en el lenguaje resultante de las mismas. En cambio, Vigotsky y sus seguidores han abordado el papel que juegan la cultura, la relación con el experto, el pensamiento y el lenguaje en la naturaleza de la conciencia humana en la interacción social.

#### **1.1.4. Vigotsky y el experto en la interacción social**

Luego de años de investigación centrados en el conflicto sociocognoscitivo, otros investigadores como Vigotsky, en su Teoría Sociocultural del Aprendizaje, plantea el estudio de la regulación interpsicológica, la intersubjetividad y la interacción como explicación de los cambios presentados en situaciones de colaboración para la solución de problemas (Martí, 1993).

Vigotsky (1934/1993) considera que el gran mecanismo para el desarrollo es la interacción social porque “La dirección del desarrollo del pensamiento no va del individual al socializado, sino del social al individual” (p. 43). El papel jugado por la interacción social es una regulación interpsicológica entre los participantes que se convierte progresivamente en autorregulación individual, a la vez, que esa autorregulación se produce en la relación entre el aprendiz y el profesor, siendo el primero de ellos no sólo capaz de aprender conocimientos, sino de autorregular su

conducta basado en un proceso psicológico interno de transformaciones (Torrego y Negro, 2012; Slavin, 2012).

La socialización está condicionada por procesos históricos, culturales y sociales que son mucho más importantes que los biológicos, puesto que el desarrollo de la psique depende directamente del entorno social en el que se desenvuelve el sujeto, siendo la sociedad la que llevará a ese aprendizaje social en el que el individuo es factor y actor fundamental (Ovejero, 1990a; Slavin, 2012; Torrego y Negro, 2012).

El lenguaje y la acción entendida como una sola acción psicológica cumplen una función primordial, orientada a conseguir la solución de un problema en un grupo. Cuanto más compleja es la actividad, o más indirecta es la meta, el papel del habla es más determinante en la interacción porque “A pesar de ser un medio expresivo y de relajar la tensión se convierte en un instrumento del pensamiento en sentido estricto, en la búsqueda y planteamiento de la solución de un problema” (Vigotsky, 1934/1993, p. 40).

De esa manera, las señales y los símbolos son considerados por Vigotsky (1934/1979) como herramientas que cumplen la función de mediadores del avance de los procesos psicológicos en donde el lenguaje es el instrumento mediador por excelencia en la interacción entre las personas con su medio ambiente.

El desarrollo individual se genera en situaciones de interacción, en la que uno de los sujetos, tiene mejor dominio de los instrumentos o herramientas mediadoras (otro compañero o profesor), siendo la relación uno-a-uno, un vínculo de mucho provecho en el legado de conocimientos para el aprendizaje. En el plano del aula, la relación directa uno-a-uno entre alumno y profesor es prácticamente imposible siendo la relación alumno-alumno, entre iguales, una posibilidad de interacción y aprendizaje mucho más alcanzable (Slavin, 2012; Torrego y Negro, 2012) “Los alumnos pueden ser mejores mediadores al tener mayor facilidad para la representación de los objetivos y de la ayuda mutua” (Torrego y Negro, 2012, p. 55).

La mediación social está íntimamente asociada con la interacción social entre el adulto o alumno experto y el aprendiz. Son precisamente los expertos, quienes favorecen la internalización de las funciones psicológicas nuevas. Esta relación entre alumno y adulto experto es lo que Bruner (1998) denominó andamiaje “...el adulto le proporciona un préstamo de conciencia al niño hasta que él desarrolla la suya propia” (p. 207). El



aprendiz va construyendo su aprendizaje como quien hace un andamio, es decir, parte por parte gracias al conjunto de elementos y fases adecuados al nivel de quien aprende (Torrego y Negro, 2012). En esa concepción tan cercana entre desarrollo y aprendizaje es donde Vigotsky formula la noción de la zona de desarrollo próximo (ZDP) que la define como

la distancia entre el nivel de desarrollo, determinado por la capacidad para resolver independientemente un problema, y el nivel de desarrollo potencial, determinado a través de la resolución de un problema bajo la guía de un adulto o en colaboración con otro compañero más eficaz (Vigotsky, 1979 citado por Martínez, Cibanal y Pérez, 2010, p. 71).

Cuando los estudiantes trabajan juntos, muchos de ellos tienen conocimientos superiores o inferiores al tema de que se trate, pero será esta ZDP la que permita que cada uno adopte el conocimiento que necesita en su propio desarrollo (Slavin, 2012). Es en las situaciones de colaboración donde se trabajan las habilidades de razonamiento mas complejas necesarias para formular tanto el problema como su solución (Vigostky (1934/1979).

Las implicaciones de esta teoría decantan en que el lenguaje no es tan sólo un medio por el cual los individuos pueden formular ideas y comunicarlas, sino que también es un recurso que hace posible que las personas piensen y aprendan conjuntamente en grupo (Mercer, 1997). Es la colaboración con un par mas competente (asimetría en el conocimiento y no en las relaciones sociales) fundamental para el desarrollo cognitivo (Hogan & Tudge, 1999). Por consiguiente, si la interacción social facilita el desarrollo cognoscitivo al ofrecer oportunidades para la discusión, comunicarse con éxito con los otros a través del diálogo implicará negociar las diferentes alternativas para alcanzar el objetivo de la tarea.

#### **1.1.5. Teoría de la Interdependencia Social**

A inicios del siglo XX, uno de los fundadores de la teoría psicológica de la Gestalt, Kurt Kufka, propuso que los grupos son dinámicas complejas en donde la interdependencia entre sus miembros puede variar.

Lewin (1935, 1948 citado por Johnson, Johnson & Smith, 2007; Johnson & Johnson, 2009b) colega de Kufka refinó su teoría e indicó que: (a) la esencia de la interdependencia entre los miembros de un grupo crea metas comunes porque el resultado es un todo dinámico que permuta de un estado al otro si alguno de los

miembros cambia de condición; (b) que en un estado de intrínseca tensión los miembros de un grupo motivan el movimiento hacia el logro de metas comunes.

En ese sentido Lewin (1935, 1948 citado por Johnson & Johnson, 2009b) enunció que la esencia de un grupo interdependiente es el resultado de que todos sus miembros vivan la dinámica grupal como un todo, en el que el cambio de cualquiera de los miembros afecta el estado de los demás por compartir las metas y las tensiones propias para alcanzarlas.

En consecuencia, y sobre la base de estos estudios, uno de los discípulos de Lewin, Deutsch (1949), extendió esta noción de las relaciones entre los miembros de un grupo para alcanzar metas comunes a partir de dos o más individuos. Así, Kufka, Lewin y Deutsch, junto con los aportes de David Johnson y Roger Johnson de la Universidad de Minnesota, dieron sentido práctico al desarrollo de los procedimientos necesarios para que los profesores pusieran en práctica dichos principios generando la llamada Teoría de la Interdependencia Social (Johnson & Johnson, 2009b).

Y ¿qué incorpora la Teoría de la Interdependencia Social de los hermanos Johnson y Johnson a las propuestas de Piaget, los neopiagetianos y Vigostky?

En esta teoría se demuestra que la *interdependencia social* ocurre cuando el rendimiento individual se ve afectado por las propias acciones del sujeto y las acciones de los otros produciéndose de esta forma una: *interdependencia social positiva* (cooperación), que se define cuando las acciones de los individuos promueven los logros de otros sujetos; una *interdependencia social negativa* (competencia), cuando las acciones de los individuos obstruyen el proceso de aprendizaje del otro (Deutsch, 1949; Johnson & Johnson, 2009b; Johnson, Johnson & Smith, 2007; Torrego y Negro, 2012) y una *no interdependencia* caracterizada por los esfuerzos individuales.

En la *interdependencia social positiva* se produce una correlación positiva entre las metas de los individuos porque estos perciben que son capaces de compartir sus metas. Este tipo de interdependencia crea procesos psicológicos asociados como son: (a) la sustitución, el grado en que las acciones de una persona sustituyen a las de otra, (b) la disponibilidad de influir en los otros y a la vez, dejarse influenciar por esos otros y, (c) la catarsis positiva, que es la inversión en energía psicológica positiva en aquellos objetos fuera de uno mismo (Deutsch, 1949).

El interés personal se expande para encontrarse con el interés de otros y las nuevas metas y motivos para alcanzarlas se crean en cooperación. Lo individual se difunde en función del interés mutuo hacia las acciones de otras personas sustituyéndolas por las propias con una inversión emocional para alcanzar las metas. Este beneficio grupal así como personal se generaliza a relaciones afectuosas y comprometidas con aquellos que trabajan por propósitos y metas similares. Demostrar esa transición desde el interés personal hacia el colectivo es, probablemente, uno de los aspectos más importantes de la Teoría de la Interdependencia Social (Johnson, Johnson & Smith, 2007).

Caso contrario, cuando se produce la *interdependencia negativa* es porque no existe una correlación positiva o alguno de los miembros del grupo no es capaz de compartir sus metas, o no son las mismas, que las del resto del grupo. El proceso que opera es la sustitución, que se puede expresar en que las acciones de una persona no sustituyen a las acciones de otra, existe resistencia a ser influido por otros y a una consecuencia que se resume en una catarsis negativa. En síntesis, el interés personal es mucho más fuerte y la motivación está orientada a ganar y a evitar la pérdida de esa fuerza (Johnson, Johnson & Smith, 2007).

Y habrá *no interdependencia* (esfuerzos individuales) cuando no exista ninguna correlación entre las metas de los individuos de un grupo. Esta situación separa a las personas, porque crea resistencia y la catarsis solo opera en el ámbito individual, lo que se traduce en que el interés y la motivación para alcanzar el logro personal se mantienen pero no se enfoca hacia el grupo, sino hacia la propia persona (Johnson, Johnson & Smith, 2007).

La premisa básica de esta teoría se sustenta en cómo la estructuración de las metas determinan las vías de interacción entre los participantes de un grupo y cómo los patrones de interacción, condicionan los logros de aprendizaje.

La Teoría de la Interdependencia Social de los hermanos Johnson es explícita al indicar que la manera en que los individuos interactúen en un grupo determinará los resultados. Promueve el aprendizaje cooperativo, entendido como interdependencia social positiva porque incrementa la motivación al logro, fortalece las relaciones interpersonales, conjuga la responsabilidad individual con la responsabilidad grupal y favorece el desarrollo de destrezas sociales (Laboratorio de innovación educativa, 2009).

### 1.1.5. Teoría de la Controversia Constructiva

¿Cuál es el aporte de esta teoría con respecto a la Teoría de la Interdependencia Social? ¿En qué se enfocaron los hermanos Johnson para dar un paso más allá de la interdependencia social positiva? La respuesta se resume en la siguiente conclusión “intellectual conflict is not only highly desirable but also an essential instructional tool that energizes student efforts to learn” (Johnson & Johnson, 2009a, p. 37).

El *conflicto* es constructivo cuando funciona como una “chispa” que brinda energía a los estudiantes a descubrir nueva información y a estudiar de manera más prolongada y con más fuerza. En grupos cooperativos, cuando se gestiona constructivamente, mejora la efectividad de los esfuerzos cooperativos (Johnson & Johnson, 2009a).

Existen tres tipos de conflictos que ocurren con frecuencia y regularidad entre los grupos: (a) *controversia constructiva*, la cual sucede cuando las ideas, información, conclusión, teoría u opinión de una persona es incompatible con la de otra pero ambas tratan de llegar a un acuerdo, (b) *conflicto conceptual*, cuando ideas incompatibles se presentan simultáneamente en el pensamiento de una misma persona o cuando la información que se recibe no es consistente con lo que ya se conoce y (c) *conflicto de intereses*, cuando las acciones de una persona que trata de alcanzar sus objetivos se ven limitadas o interferidas por las acciones de otra persona para alcanzar sus propias metas.

Entonces ¿qué tipos de conflictos se ven implicados en el conflicto cognitivo? La controversia constructiva y el conflicto conceptual (Johnson, Johnson & Smith, 2007) y, si la *controversia constructiva* implica que aunque dos personas disientan ambas intentan llegar a un acuerdo ¿cómo opera y se gestiona esa discrepancia positiva?

Estas dos preguntas encuentran respuesta cuando surge la controversia en un grupo y puede ser gestionada con criterios constructivos (cooperación) o destructivos (competencia) dependiendo de los niveles de relación y de manejo de las herramientas de los subgrupos o del grupo. En el trabajo constructivo, la controversia promueve la incertidumbre acerca de la forma correcta de enfocar un punto de vista, genera la búsqueda activa de información, reconceptualiza el conocimiento individual y las conclusiones grupales, incrementa el dominio y retención del material usado en la discusión y favorece un mayor juicio sobre el razonamiento adoptado de un tópico. En cambio, los individuos que trabajan solitariamente en situaciones competitivas e individualistas no tienen la oportunidad de experimentar tal proceso y en consecuencia,

sufre la capacidad productiva y la calidad de las decisiones adoptadas (Johnson & Johnson, 2009a).

Las controversias tienden a ser constructivas cuando el contexto de la situación es cooperativo, los miembros del grupo son heterogéneos, la información y la experiencia es distribuida entre los grupos, los miembros cuentan con suficientes herramientas para resolver los conflictos y los cánones de la argumentación son respetados (Johnson, Johnson & Smith, 2007).

### *Acuerdos y desacuerdos entre las teorías*

Si se ha mencionado en todo este apartado de las bondades de cada una de las teorías psicopedagógicas que convalidan las situaciones de aprendizaje cooperativo, en donde se afirma que la cooperación ofrece mejores resultados académicos en comparación con situaciones de aprendizaje competitivas e individuales, no es menos cierto que existen diferencias entre ellas que hasta los momentos son difíciles de reconciliar y que a la vez requieren seguir trabajándose desde un marco teórico y metodológico.

Son los mismos hermanos Johnson los que reconocen esta carencia y así lo expresan

La teoría de la interdependencia social supone, por ejemplo, que los esfuerzos cooperativos se basan en la motivación intrínseca generada por factores interpersonales al trabajar juntos y en aspiraciones conjuntas para alcanzar algún objetivo significativo; mientras que la teoría conductista sostiene, por su parte, que los esfuerzos cooperativos dependen de la motivación extrínseca para obtener recompensas. La teoría de la interdependencia social está constituida por conceptos relacionales que se ocupan de lo que sucede entre las personas (es decir, la cooperación es algo que existe sólo entre las personas, no dentro de cada una de ellas), en tanto que, por el contrario, la perspectiva evolutiva cognitiva se centra en lo que ocurre dentro de cada persona (es decir, el desequilibrio, la reorganización cognitiva) (Johnson y Johnson, 1999c, p. 264).

La necesidad de seguir investigando en los límites, a veces difusos entre estas teorías y en ocasiones contradictorios, continua siendo de interés para quienes apuestan por el aprendizaje cooperativo como metodología de enseñanza-aprendizaje.

## ***1.2. Aprendizaje individual, competitivo y cooperativo***

¿Dónde está el límite entre el aprendizaje cooperativo, competitivo e individual? ¿En qué se diferencian cada uno de estos aprendizajes? ¿En qué medida se favorece el aprendizaje cooperativo, competitivo e individual en la programación universitaria?

La línea que divide al aprendizaje según los objetivos que se pretendan alcanzar resulta en ocasiones no tan clara o “controversial”. Cada uno de estos tres tipos de aprendizajes

ha tenido a lo largo de los años seguidores y detractores. El contexto social y cultural, la visión de hombre desde la filosofía y la sociología, las corrientes pedagógicas y la teoría psicológica que sustenta cada aprendizaje marcaron diferencias a favor y en contra. Una revisión de la historia del aprendizaje así lo demuestra (Johnson & Johnson, 1999b).

Entre los años 30 y finales de los 70, la educación abogaba por el *aprendizaje competitivo*. Los resultados obtenidos en cada clase se evaluaban en función de la curva de la normalidad. Por consiguiente, el desempeño de cada alumno estaba “marcado” por su posición en la curva lo que indicaba el resumen de su rendimiento. Skinner (1954) afirmaba que el aprendizaje individual era beneficioso para la mejora del rendimiento pero que la “Cooperation can also be set up, perhaps more easily than competition” (p. 48).

Algunos investigadores que han estudiado este tipo de aprendizaje, muchos de ellos a favor de este tipo de situaciones como Deutsch (1949) de quien se ha usado con mayor frecuencia la definición de competencia entendida como la “correlación negativa entre los objetivos de los diferentes participantes en una situación dada” (Johnson y Johnson, 1999c, p. 185); Michaels (1977) quien apostó por este tipo de aprendizaje alegando que sus logros superaban el aprendizaje cooperativo; Skinner (1968) quien desde la teoría conductista la definió como la consecuencia de una recompensa a aquel que ha logrado los resultados más altos en comparación con los demás; Kelley & Thibaut (1978), colegas conductistas de Skinner, quienes colocaron el acento en la forma de optimizar las propias recompensas y minimizar los costos en relación con los demás; y Helmreich, Beane, Lucker & Spence (1978) para quienes la competencia es la ganancia en situaciones interpersonales (véase a todos estos autores en Johnson & Johnson, 1999b).

Sin embargo, otros como Kohn (1992 citado por Johnson & Johnson, 1999b), señalaron sus efectos negativos por tratarse de un aprendizaje que incrementa la ansiedad, promueve la motivación extrínseca y favorece la relaciones negativas entre compañeros y profesores.

Seguidores y detractores también acompañaron al *aprendizaje individual* entre los años 70 y 80. Hasta el mismo Slavin (1977) así lo afirmaba en sus inicios, en el artículo *Classroom reward structure: an analytical and practical review*, asegurando que los resultados eran mucho más beneficios en comparación con el aprendizaje cooperativo “unless subjects have important resources to share or withhold at their discretion,

competitive and individual reward structures are more effective than cooperative ones for increasing performance” (p. 647).

En cambio, Bronfenbrenner (1974) criticó a quienes abogaban por este tipo de aprendizaje asegurando que el aprendizaje no solo individual sino también el competitivo promovían la alienación y el aislamiento porque las personas se sienten hostiles y desconectadas de las actividades de su entorno.

La década de los 70 estuvo marcada por un gran número de estudios que se centraron en comprender las diferencias entre el aprendizaje cooperativo, el aprendizaje competitivo y el aprendizaje individual. Slavin y los hermanos Johnson formaron a profesores universitarios en el *aprendizaje cooperativo* y así lograron corroborar los beneficios de este aprendizaje en el rendimiento académico en comparación con los otros dos métodos tradicionales (Johnson & Johnson, 1999b).

En el aprendizaje individual, no existe una interdependencia para el logro de los objetivos porque el desempeño de cada individuo no se relaciona con el rendimiento de los demás (Johnson & Johnson, 1999b).

En contraste, en el aprendizaje competitivo si hay interdependencia entre los objetivos; pero, es negativa porque para alcanzar el logro personal se requiere que los otros no alcancen los propios. Los valores son opuestos al éxito académico. Las diferencias entre “quienes saben más” y “quienes saben menos” son cada vez mayores. Los estudiantes con dificultades de aprendizaje encuentran un mayor impedimento en este tipo de aprendizaje por recibir con frecuencia una retroalimentación negativa por sus esfuerzos, siendo necesario encontrar otra manera de alcanzar imágenes positivas de sí mismos (Slavin, 1999a).

En el aprendizaje cooperativo, la interdependencia es positiva para el cumplimiento de las metas. La interacción entre alumno–alumno es dependiente porque para alcanzar los objetivos individuales se requiere que los otros miembros del grupo también logren los suyos

La idea subyacente es que si los alumnos quieren tener éxito como equipo, estimularán a sus compañeros para que luzcan y los ayudarán a hacerlo. Con frecuencia los alumnos pueden explicar las ideas muy difíciles a sus compañeros, ya que saben traducir el discurso del docente a su propio lenguaje (Slavin, 1999a, p. 18).

Sin embargo, todos los equipos de trabajo pasan por problemas similares: las metas individuales deben armonizarse con las colectivas porque (a) las estructuras de recompensa tienen que ser visualizadas como posibles por los miembros del grupo, y (b) la responsabilidad y la interdisciplinariedad deben equipararse para que ninguna pese más que la otra. Es el establecimiento de un lenguaje común lo que permitirá la cooperación y la colaboración de modo transversal en la estructura organizativa del propio grupo (Foyle & Shafto, 1995).

La tabla comparativa 1, en ese esfuerzo por diferenciar los tres tipos de aprendizaje, expone de cada uno de ellos su definición, los objetivos que pretenden alcanzar, cómo percibe el estudiante el aprendizaje, los mensajes explícitos e implícitos inmersos en cada situación de aprendizaje, los niveles o no de cooperación, cómo ocurre la interacción entre estudiantes y por último, la evaluación de los resultados:



Tabla 1: Diferencias entre aprendizaje individual, cooperativo y competitivo

Aspectos	Aprendizaje individual	Aprendizaje competitivo	Aprendizaje cooperativo
Definición	El trabajo es individual para alcanzar el objetivo de tareas que no implican relación con los otros. Las tareas no requieren vínculos e interdependencia con otros estudiantes. El esfuerzo se centra en el trabajo individual para que el propio aprendizaje alcance la meta establecida independientemente del esfuerzo de los demás	El esfuerzo se centra en el desempeño del alumno en comparación con sus compañeros. Para alcanzar el objetivo se “trabaja en contra” de los demás. Un alumno o solo unos pocos logran alcanzar la meta	Trabajar juntos para alcanzar metas compartidas. El beneficio es individual y grupal. Los estudiantes trabajan juntos para mejorar su propio aprendizaje y el de los demás
Objetivos	El alcanzar los objetivos no influye en los demás. Es independiente de los compañeros	Se promueve la capacidad de rapidez y de precisión para que el desempeño sea mejor que el de los compañeros	Por lo general, se trabaja en grupos cooperativos heterogéneos en donde se aprenden los materiales asignados y se asegura que los integrantes del grupo hagan lo mismo
Percepción del estudiante	En situaciones muy individualistas, los estudiantes buscan beneficios personales	Se alcanzan los objetivos solo si los otros fracasan en alcanzar los propios. Los estudiantes se esfuerzan por ser mejores que sus compañeros. Trabajan para desplazar a los demás. El alcance de la meta es solo para algunos. Mayor brecha entre quienes “saben” y “quienes no saben”	Todos los miembros del grupo comparten una meta común, se trabaja para el “bien” colectivo, el desempeño individual obedece al esfuerzo propio y también al de los compañeros. El éxito se celebra desde lo individual pero también desde lo colectivo
Mensajes explícitos/ implícitos	Cuán bien puedo desempeñarme, qué ganaré con ello, si soy capaz, recibiré una buena calificación, lo logré, si mis compañeros triunfan o fracasan, no me afectará	Cuánto más ganes, peor para mí; cuanto más gane yo, peor para ti; puedo vencerte; mi ganancia significa tu pérdida; sólo unos pocos podrán sacarse un 10; tu fracaso facilita mi éxito; los ganadores siempre ganan, los perdedores siempre pierden	“Nos salvamos juntos o nos hundimos juntos”; “Tus esfuerzos me benefician y mis esfuerzos te benefician”; “La unión hace la fuerza”; “Juntos, podremos lograr lo que nos proponíamos”
Niveles de cooperación		Centrados en el grupo (el mejor del grupo), en la clase (el mejor de la clase), en la escuela (el mejor de toda la escuela) o en el país (buscar desempeñarse mejor que otro ciudadano)	En toda la clase, al asegurarse que todos hayan aprendido el material que ha sido asignado y en la escuela, porque se busca el progreso de todos los estudiantes que conforman la institución educativa
Esquema de interacción		Los estudiantes dificultan e interfieren el éxito de los demás. Prevalece el trabajo individual. Se oculta información. Se rehúsa la ayuda del otro	Los estudiantes promueven el éxito de los demás. Discuten, explican, escuchan y se ayudan entre sí. Esta interacción ocurre tanto entre los grupos como

			dentro de los grupos.
Evaluación de resultados		Fundamentado en normas. El desempeño se ordena en función del mayor al menor rendimiento	Centrado en el aprendizaje y progreso de cada estudiante pero también en el grupo de clase, el aula y la escuela

Fuente: adaptación de Johnson y Johnson (1999c, pp. 21–24)

Johnson y Johnson (1999b) señalan que no solo es posible sino que es indispensable la integración en clase del aprendizaje individual, competitivo y cooperativo pero para estos autores, el peso de la programación recae en objetivos y actividades que favorezcan el aprendizaje cooperativo.

Sugieren que entre el 60 y 70 por ciento del tiempo de clase debería estar diseñado para alcanzar este tipo de aprendizaje. La competencia entre estudiantes puede utilizarse luego de una actividad cooperativa, como un cambio de dinámica que proporciona un rato distendido entre los alumnos y, el aprendizaje individual preceder a una actividad cooperativa. Lo deseable es que las actividades siempre empiecen y terminen con esfuerzos cooperativos.

Si se trabaja con los estudiantes en un ambiente cooperativo donde se fomenta la ayuda y la colaboración entre ellos, si se conocen entre sí y trabajan juntos para un beneficio común, se sentarán las bases para situaciones constructivas tanto competitivas como individualistas, en donde se compite por placer y lo individual beneficia el progreso propio y colectivo, minimizando así las dificultades propias de cada metodología (Johnson & Johnson, 1999b).

Dos máximas acompañan las unidades integradas de aprendizaje “el aprendizaje individualista y el competitivo no pueden alcanzar su máximo potencial al menos que se los use en un contexto fuertemente cooperativo” (Johnson y Johnson, 1999c, p. 238) y “Todo aprendizaje cooperativo es aprendizaje en grupo, pero no a la inversa, pues el aprendizaje en pequeño grupo puede ser también de carácter individual, competitivo o no” (Nieto, 2004, p. 105). La competencia constructiva se nutre del reto individual, en donde cada estudiante compite consigo mismo y con el grupo, superación que contribuye a una competencia constructiva porque la meta común no desaparece: el para qué es social y se construye desde lo individual.

Son Johnson, Johnson y Smith (1998) los primeros en afirmar que el aprendizaje cooperativo debe retornar a la universidad y romper con el mito del aprendizaje individual como la metodología que prevalece porque “Despite the remarkable

achievements of academic teams, the myth of the genius individual still exists; it underlies educational practice that assumes each student should work separately and apart from classmates” (p. 27).

### ***1.3. Aprendizaje cooperativo y aprendizaje colaborativo***

Los términos aprendizaje cooperativo, aprendizaje colaborativo, trabajo en grupo o trabajo en equipo se emplean indiferentemente debido a que en el uso cotidiano, colaborar y cooperar guardan un sentido similar, disyuntiva que en la literatura científica no está exenta de polémica (Barkley, Cross & Howell, 2007; Markulis & Strang, 2002).

Desde la etimología, colaborar proviene del latín *co-laborare*, *laborare cum* y significa trabajar juntamente con y cooperar, del latín *co-operare*, *operare cum*, refiere a trabajo conjunto y además, ayuda, interés, servicio y apoyo (Corominas y Pascual, 2007). Así, cooperar amplía su significado hacia ayuda, apoyo mutuo e interesarse por el otro.

#### ***El Aprendizaje Colaborativo***

Definen Damon y Phelps (1989) a la colaboración como una relación cuyo propósito es la obtención o aplicación de un conocimiento en un grupo de dos o más estudiantes que posean habilidades análogas. Es decir, es una relación de elevada colaboración simétrica y mutualidad (Duran y Monereo, 2012) en donde es necesario que exista: (a) simetría en la acción, entendida como el mismo rango de interacciones permitidas por cada agente de aprendizaje; (b) simetría del conocimiento o herramientas de conocimiento, lo que no se debe confundir con puntos de vista convergentes, sino capacidades similares; y, (c) simetría de estatus, que es relativa al grupo social al que pertenecen los miembros del grupo (Dillenbourg, 2003).

Bruffee (1999) y Cooper (2009), autores que se dedican al estudio del aprendizaje colaborativo en el ámbito universitario, desarrollan el tema. El aprendizaje colaborativo se caracteriza por ser poco estructurado, en ocasiones hasta informal, en donde la responsabilidad recae en los individuos que conforman el grupo. Es el aprendiz quien diseña y controla todas las decisiones que giran en torno a su aprendizaje.

El profesor de este tipo de grupos colaborativos tiende a confiar a sus estudiantes que se autogobiernen en contextos sustantivos. Este énfasis en el autogobierno es, en sí mismo, una de las grandes metas del aprendizaje colaborativo: ayudar a los jóvenes estudiantes

a reconocer disensos y desacuerdos, y a que cooperen los unos con los otros en diferenciarlos (Bruffee, 1999).

De esa manera el aula de clase para el aprendizaje colaborativo cambia de estructura con respecto al aula tradicional, porque los estudiantes buscan negociar relaciones fundadas sobre lo pequeño, en su comunidad de aprendizaje, con otras comunidades de aprendizaje y a la vez, con el profesor. Es una relación de interdependencia y esta distribución del aula ayuda a los alumnos a ser autónomos, a articularse y a ser sociales e intelectualmente más maduros. Los grupos colaborativos favorecen un aprendizaje de los tópicos sustantivos y a no concluir sobre elementos fácticos sino a construir sobre el resultado de una disciplina social producto de un proceso de conversación, investigación y negociación (Bruffee, 1999).

El aprendizaje colaborativo significa que los estudiantes trabajan en conjunto, en grupos pequeños o en parejas con el objetivo de obtener un aprendizaje común. Es una fórmula que suma estudiantes para que no trabajen en solitario. Barkley, Cross y Howell (2007) distinguen tres características de este tipo de aprendizaje. La primera, el diseño educativo es intencional, es decir, los profesores definen a priori la actividad colaborativa, el resultado que se pretende alcanzar, las vías para alcanzarlo y saben cuáles pueden ser los tropiezos de sus alumnos en la experiencia. La segunda característica, los alumnos trabajan activamente con actitud colaboradora. Sin ser suficientes las dos características anteriores, se suma la tercera que refiere a un resultado positivo de una experiencia significativa, es decir, incrementa los conocimientos y ahonda en los contenidos de la asignatura que se estudia.

### *El Aprendizaje Cooperativo*

El aprendizaje cooperativo tiene, además del nombre propio, un par de denominaciones más: aprendizaje entre iguales o aprendizaje entre colegas. Lo destacable es que cumple con la máxima de que el mejor maestro de un estudiante es otro estudiante (Ferreiro, 2006, 2010). Sobre esa base, la cooperación como característica de la humanidad es la forma de interacción en la vida diaria, en empresas y comunidades (Olmos, 2012).

El aprendizaje cooperativo emplea en la enseñanza a pequeños grupos de estudiantes que trabajan juntos en potenciar su aprendizaje (Smith, 1996 citado por Barkley, Cross & Howell, 2007). El profesor adopta dos roles: el de experto de su asignatura y el de ser la autoridad en el salón de clases. Así, preparar las clases, asignar las tareas de grupo,

controlar el tiempo, diseñar los materiales, supervisar el aprendizaje de los alumnos y velar porque todos trabajen cooperativamente en la labor es parte de la función del profesor (Cranton, 1996; Smith, 1996, citados por Barkley, Blosser, 1992, Cross & Howell, 2007 citado por Olmos, 2012; Fathman & Kessler, 1993, citados por Trujillo, 2006).

Evidentemente este tipo de enseñanza comporta que el profesor es experto en su área de conocimiento, por lo que su juicio sobre lo ‘correcto o incorrecto’ es asumido como adecuado, razón por la cual las experiencias de estos docentes en sus aulas, al llevarlas a una expresión de conocimiento empírico, muestran mayor satisfacción por los resultados de aprendizaje en esta modalidad que en el aprendizaje individual (Astin, 1993; Light, 2001; Pascarella & Terenzini, 1991 citados por Barkley, Cross & Howell, 2007).

Los grupos cooperativos se caracterizan por ser de tamaño reducido y heterogéneamente compuestos en su rendimiento académico. La participación es equitativa de manera que todos los alumnos puedan ser parte de la tarea y, en la medida de lo posible, que esta participación sea simultánea para aprender contenidos aprovechando al máximo esas potencialidades. Estas condiciones preliminares hacen que se creen una serie de responsabilidades que bien define Pujolàs (2008):

- Los miembros del grupo o equipo de trabajo tienen dos responsabilidades: aprender ellos mismos las enseñanzas del maestro y, por otro lado, contribuir en el aprendizaje de sus compañeros (también Balkcom, 1992 citado por Trujillo, 2006).
- Se busca que el alumnado esté organizado en equipos de trabajo para el aprendizaje conjunto.
- El profesor usa el aprendizaje con dos fines: para que el alumno aprenda los contenidos y a trabajar en equipo, siendo esta una competencia más dentro del plan de estudios.
- Y, se suma a esta lista, el que el alumno deje de depender del profesor y se fomente un pensamiento autónomo y reflexivo (Olmos, 2012).

Las estructuras cooperativas de aprendizaje contienen un entramado de recompensas y esfuerzos individuales que son proporcionales a la calidad del trabajo en grupo (Kelly & Thibaut 1969, citados por León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

Slavin (1985) apunta que el aprendizaje cooperativo es un método que engloba el incentivo a la cooperación, es decir, otorga a los estudiantes reconocimientos, recompensas y calificaciones sobre el esfuerzo común.

Entonces ¿son similares o difieren los términos aprendizaje cooperativo y aprendizaje colaborativo? ¿se pueden utilizar indistintamente en el ámbito universitario? ¿cómo los ha abordado la literatura científica?

El aprendizaje cooperativo y el aprendizaje colaborativo comparten en común que existe un grupo con el que se trabaja, en donde se respeta el conocimiento de cada estudiante, se apuesta por ese conocimiento como potencial de éxito y, el origen de procedencia de ambos, es el constructivismo. Pero, a la vez se diferencian en cuanto a su filosofía, estructura, la relación estudiante–profesor y el fin que pretenden conseguir.

Así, cooperar amplía su significado hacia ayudar a, secundar al otro siendo mucho más estructurado y más enfocado que el trabajo colaborativo (Millis & Cottell, 1998). En esta línea, Bruffee (1995) señala que la meta del aprendizaje cooperativo es trabajar juntos en armonía y en apoyo mutuo para hallar la solución, en cambio, el aprendizaje colaborativo busca desarrollar personas autónomas y reflexivas. Este autor defiende que la expresión *cooperativa* es propia del trabajo con niños o adolescentes y *colaborativa* es más pertinente para universitarios o adultos, a lo que añade

The major disadvantage of collaborative learning is that in nurturing the educational rewards to be gained from self-governed student peer relations, it sacrifices guaranteed accountability. The major disadvantage of cooperative learning is that in guaranteeing accountability, it risks maintaining authority relations within each small working group and in the class as a whole that replicate the authority relations of traditional education (Bruffee, 1995, p. 18).

Por su parte, Panitz (1999) considera que la colaboración es una filosofía de la interacción y un estilo de vida personal donde los individuos son responsables de sus acciones, incluyendo el aprendizaje y el respeto de las capacidades y aportaciones de sus compañeros, mientras que la cooperación, es una estructura de interacción diseñada para facilitar el logro de un producto final específico o meta a través de las personas que trabajan juntas en grupos.

El aprendizaje colaborativo no es sólo una técnica de aula sino una manera de vivir y tratar con otras personas. Es el aprendiz quien diseña y controla todas las decisiones que giran en torno a su aprendizaje. En cambio, en el aprendizaje cooperativo la situación de aprendizaje es estructurada de tal forma que todos los miembros del grupo buscan

alcanzar una meta en común, convirtiendo la experiencia de aprendizaje en un ejercicio mucho más profundo que una mera forma de agrupamiento entre estudiantes. Es el profesor, y no el aprendiz, sobre quien recae la estructura y división de las funciones del grupo, las cuales se orientan al cumplimiento de un objetivo (Bruffee, 1999; Cohen, 1994b).

Cooper (2009) y sus colegas Matthews, Davidson & Hawkes (2009) hacen notar que el aprendizaje colaborativo se diferencia sustancialmente del cooperativo en: (a) la relación del profesor con los alumnos en cuanto a estilo, función y grado de involucramiento; (b) la autoridad y las relaciones de poder entre alumno y profesor; (c) el tiempo necesario para adaptarse al trabajo en equipo; (d) las formas en cómo el conocimiento es asimilado; (e) el propósito de los grupos al enfatizar sus dominios, desarrollo de juicios y construcción del conocimiento; (f) la importancia de los aspectos personales, sociales y cognitivos entre los alumnos y (g) la variada implementación, desarrollo de tareas y el grado de distribución del tiempo y asignaciones para alcanzar la meta.

Es la comunidad científica la primera en dar el paso creando puentes entre ambas acepciones porque

Followers of the two traditions have published in different journals, created bibliographies with few common names, sponsored different conferences, and for many years, had little contact with each other. As a result, among collaborative –and cooperative– learning practitioners there exists a certain amount of ignorance about the other tradition, an about variations within each school of thought. Lack of information and lack of communication among proponents of different viewpoints have sometimes led to a rigid sense of orthodoxy (Matthews, Cooper, Davidson & Hawkes, 2009, pp. 11–12).

Más allá de estas diferencias, es cierto que en la literatura científica el aprendizaje cooperativo, en comparación con el aprendizaje colaborativo, es el que mejor se ha definido y trabajado en todos los niveles educativos, cuenta con un número importante de documentación empírica que lo respalda y ha logrado una mejor operacionalización de la metodología (Cuseo, 1992).

Como cooperar implica un paso mas allá de la colaboración, es fundamental que se continúe investigando sobre la cooperación porque incidirá positivamente en las múltiples mediciones de resultados sin que ello menosprecie, en palabras de Millis y Cottell (1998), el subrayar las dos aproximaciones –colaborativa y cooperativa– con

miras al logro académico.

#### ***1.4. Elementos esenciales del aprendizaje cooperativo***

¿Qué principios o elementos son esenciales para un efectivo aprendizaje cooperativo? El aprendizaje cooperativo se caracteriza por una serie de factores o variables que afectan la relación de aprendizaje en grupo. Son componentes que parten, fundamentalmente, de la existencia de un grupo cuyo valor principal es compartir como meta el adquirir conocimiento. Esta condición es la que permite afirmar que existirá interacción en el seno de ese grupo.

Como en todo proceso evolutivo de la ciencia, las relaciones y conceptualizaciones no se paralizan. Al contrario, se dinamizan y se contrastan de tal forma que diversos autores tratan de comprender los mismos fenómenos. En ese sentido, Johnson y Johnson (1999a) y Kagan y Kagan (2009, 2011) comenzaron a diseñar estructuras simples para explicar los elementos esenciales del aprendizaje cooperativo.

Para Johnson y Johnson (1994, 1999d) son cinco los elementos fundamentales para lograr un aprendizaje cooperativo: *interdependencia positiva* y *el conflicto intelectual*, *responsabilidad individual/personal*, *interacción cara a cara*, *habilidades interpersonales y en grupo* y *el procesamiento/evaluación grupal*.

Para Kagan y Kagan (2009, 2011) son cuatro los elementos básicos los que se recogen bajo el acrónimo IRPI –en inglés PIES: Positive Interdependence, Individual Accountability, Equal Participation y Simultaneous Interaction–. Así, se puede definir mejor a esos elementos fundamentales preguntando de la siguiente forma: *interdependencia positiva* ¿se encuentran los estudiantes del mismo lado? ¿la tarea requiere del trabajo de todos? *responsabilidad individual* ¿cada miembro asume su parte de responsabilidad dentro del grupo? *participación equitativa* ¿es la participación entre todos aproximadamente equitativa? e *interacción simultánea* ¿qué porcentaje de estudiantes están interactuando abiertamente a la vez?

A pesar de que a simple vista pareciera que ambos planteamientos disienten en esencia, comparten los mismos elementos fundamentales aunque bajo nombres distintos, en donde, en caso de faltar al menos uno de estos principios el trabajo en equipo no sería el resultado de un aprendizaje cooperativo.



La *interdependencia positiva* se logra cuando los estudiantes no pueden alcanzar una meta común sin la aportación de cada uno de los miembros del grupo (Johnson & Johnson, 2014a).

Se entiende como la dependencia mutua entre los miembros de un grupo y a la vez, el sentido de esa dependencia puede ser positivo, negativo o ausente. Una relación positiva significa que el éxito de un estudiante depende del éxito del otro, lo que conduce a una clase cooperativa. Una relación negativa refiere a que el éxito de un alumno disminuye la posibilidad del éxito del otro, lo que se traduce en una clase competitiva. La ausencia de relación implica que los estudiantes trabajan de manera independiente y la ganancia individual de un estudiante no se relaciona con la ganancia de los otros y en consecuencia, la clase es individual (Kagan & Kagan, 2009, 2011).

Además de la explicación de Kagan y Kagan, Johnson y Johnson (1999d), apuntan y coinciden con Deutsch (1949) que la interdependencia positiva pretende crear fuerzas responsables que se añaden al concepto de responsabilidad moral de uno con otro como parte de un grupo, de una totalidad que busca la motivación de los miembros del equipo para satisfacer las normas de los pares (Deutsch, 1949).

La interdependencia positiva obliga a todos los miembros del grupo a alcanzar el resultado que sientan desde su ámbito de responsabilidad para: (a) completar y compartir el trabajo y, (b) para facilitar el trabajo de todos los miembros del grupo. Adicionalmente, cuando una persona afecta los logros de sus colaboradores se sentirá igualmente afectado a sí mismo razón por la cual el miembro del grupo realizará “un esfuerzo mayor para que su rendimiento no sea una causa que conduzca al fracaso del grupo” (Matsui, Kakuyama & Onglatco, 1987, p. 407).

En ese sentido, sentirse que se falla es malo, pero sentir que se falla a otros, es peor. Así, una persona será agradable y respetada por los compañeros de grupo en la medida del compromiso y cumplimiento de la responsabilidad adquirida con su grupo “la combinación de interdependencia de objetivos y recursos aumenta los logros por sobre la interdependencia sola y la interdependencia de recursos no aumenta los logros a menos que también haya interdependencia de objetivos” (Johnson y Johnson, 1999c, p. 122).

Existen distintos grados de interdependencia. Un nivel fuerte de interdependencia se produce cuando para alcanzar la meta, el aporte de cada uno de los miembros del grupo

es esencial. Esta fuerza se va haciendo más débil en la medida en que la participación de cada miembro no es tan necesaria para alcanzar los objetivos comunes de aprendizaje (Kagan & Kagan, 2009, 2011) hasta llegar a un nivel tan endeble, en donde la contribución de cada uno de los miembros del equipo no podría quizás ayudar al éxito del grupo.

La *responsabilidad individual* existe cuando cada uno de los miembros del grupo requiere hacer una contribución que es conocida por el profesor y por los otros miembros del grupo. Esta responsabilidad no puede ocultarse ni evadirse (Kagan & Kagan, 2009, 2011). Se logra cuando se evidencian estos tres componentes: individual (el desempeño o rendimiento se logra sin la ayuda de los otros), público (alguien es testigo del desempeño) y requisito (el rendimiento es un requisito en si mismo).

Asimismo, la responsabilidad grupal e individual obliga a incrementarse cuando existe un grupo que tiene individuos responsables. De esa manera, la responsabilidad grupal se logrará cuando los éxitos del equipo se evalúen con los resultados de cada miembro del grupo y esta valoración se compare con el estándar que se supone se debe alcanzar. Por su parte, la responsabilidad individual se conquistará cuando los resultados individuales de cada uno de los miembros del grupo se comparen nuevamente con el estándar esperado.

Hooper, Ward, Hannafin & Clark (1989 citados por Johnson & Johnson, 2009b) hacen notar que la cooperación es el resultado del alto logro que las distintas responsabilidades individuales han estructurado para alcanzarlo o no. Agregan al concepto Archer-Kath, Johnson y Johnson (1994) que la responsabilidad grupal se fundamenta en el incremento de la responsabilidad individual y que esta es percibida como interdependiente por los miembros del grupo “las habilidades sociales es una de las variables moderadoras para la productividad del grupo” (p. 693).

La falta de responsabilidad individual podría reducir los sentimientos de responsabilidad personal. Los miembros del grupo pueden disminuir sus contribuciones hacia las metas a alcanzar cuando el grupo trabaja por objetivos en los que es difícil identificar las contribuciones de cada miembro del grupo, lo que hace que la cohesión del grupo se debilite y se diluya lo individual para alcanzar la meta final (Harkings & Petry, 1982; Ingham, Levinger, Graves & Peckham, 1974; Kerr & Bruun, 1981; Latane,

Williams & Harkins, 1979; Moede, 1927; Petty, Harkings, Williams & Latane, 1977; Williams, 1981; Harkings & Latane, 1981 citados por Johnson & Johnson, 2014a).

Sin embargo, si se produce una alta responsabilidad individual es mucho más claro cuál de los miembros del grupo es responsable hasta el final de la tarea, se evitan las reiteraciones en la actividad y, si se logra que cada uno de los miembros se haga responsable de su labor, la cohesión de grupo hará que la holgazanería se diluya (Johnson & Johnson, 2014a).

Generalmente, cuando un grupo se hace cada vez más grande, los miembros del equipo están más a disgusto con sus contribuciones personales y ven como menos importante que el grupo alcance el éxito (Kerr, 2001 citado por Johnson & Johnson, 2009b). Si un grupo incrementa su tamaño, la tendencia a la comunicación individual es menos frecuente, ya que se puede reducir la cantidad de información para la toma de decisiones “Increasing the size of the disagreeing majority resulted in increased yielding” (Gerard, Wilhelmy & Conolley, 1968, p. 79) y se hace menos fidedigna porque los miembros del grupo pueden alterar sus posiciones para adecuarse o conformar lo que consideran debe agradar al grupo (Gerard, Wilhelmy & Conolley, 1968).

La holgazanería social se incrementa en la misma medida en la que crece el grupo. A menor tamaño del grupo mejores resultados se van a alcanzar. Añaden Morgan, Coates y Rebbin (1970 citados por Johnson & Johnson, 2009b) que los alcances del grupo realmente mejoran cuando se pierde a un miembro del grupo porque probablemente los miembros restantes considerarán que sus contribuciones son mucho más necesarias.

La *participación equitativa* refiere una participación lo más “equilibrada” posible, sabiendo que es imposible asegurar una equidad al 100%. Los estudiantes sienten respeto mutuo y se benefician de un estatus de igualdad entre sus compañeros del grupo. Hablar de equidad implica participación no voluntaria. Así, los estudiantes de bajo rendimiento y con necesidades especiales son los que más se benefician porque la participación no es una opción (Kagan & Kagan, 2009, 2011). Para Johnson y Johnson (2014a) el sistema de equidad se sostiene sobre la base de una justicia distributiva porque todos los miembros del grupo reciben los beneficios de la cooperación en función de la proporción de sus contribuciones. De esta manera, las recompensas también forman parte del sistema porque aquellos miembros del grupo que mas contribuyen al éxito del equipo reciben mayores beneficios.

Si la interdependencia positiva por sí misma puede lograr incrementar el resultado del grupo, la *interacción promotora cara a cara* o la *interacción simultánea* –según el autor que se consulte– es lo que mas influye en la productividad de los grupos, en la competencia social y en la afectividad. Para Johnson y Johnson (1994, 2009a) la interdependencia positiva está definida desde la *interacción promotora cara a cara* caracterizada por una ayuda eficaz y eficiente, el intercambio de recursos e información, una retroalimentación permanente, la toma de decisiones para la comprensión de los problemas, el esfuerzo constante para alcanzar los objetivos, el trabajo sobre la confianza, la confiabilidad y la motivación.

En cambio, la interacción opuesta ocurre cuando el desaliento individual bloquea y obstruye los esfuerzos de unos y otros para alcanzar las metas; el foco de los individuos, por lo tanto, se vuelca en ser productivos y prevenidos frente a cualquier otro que lo sea más que él mismo. Y, la no interacción aparece cuando los sujetos trabajan independientemente, sin intercambio entre ellos, buscando el logro individual y siendo productivos solo para sí mismos, ignorando o considerando irrelevantes los esfuerzos de los demás (Johnson & Johnson, 2009a).

Por su parte, Kagan y Kagan (2009, 2011) matizan la propuesta de Johnson y Johnson comprendiendo la interacción promotora como la *interacción simultánea* que refiere a cómo los estudiantes se mantienen activamente ocupados todo el tiempo aumentando así el compromiso entre todos los miembros del grupo. Los alumnos participan a la vez y no en una estructura secuencial de turnos de habla, como ocurre en la clase tradicional, en donde el profesor asume el mayor tiempo de la clase y los estudiantes en contadas ocasiones. Para estos autores el número de integrantes del grupo, par o impar, incide en el número de interacciones simultáneas. En los equipos pares habrá una mayor probabilidad de que todos los miembros participen simultáneamente mientras que en un grupo impar alguno puede quedar en un momento dado fuera de la actividad.

El *procesamiento de grupo* es el último de los elementos esenciales de la cooperación y ocurre cuando cada uno de los miembros: (a) refleja acciones que tienden a ayudar o no y (b) toman decisiones acerca de qué acciones deben continuar y cuáles no. El propósito del grupo será procesar con claridad las mejoras para la efectividad entre los miembros y asumirlas como necesarias para alcanzar las metas grupales. Un grupo es efectivo cuando permanentemente reflexiona sobre sus acciones y planifica cómo mejorar el trabajo (Johnson & Johnson, 1994, 1999a, 2009b).

De esta forma, el procesamiento se entiende como una secuencia de eventos identificables en un momento del tiempo. El procesamiento de objetivos, la secuencia de eventos instrumentales para alcanzar las metas de rendimiento y el procesamiento grupal se logra cuando se cumplen los siguientes pasos: (a) evaluación de la calidad de la interacción entre los estudiantes, (b) retroalimentación del profesor a cada grupo, (c) planificación de los objetivos por parte de los grupos para mejorar su eficacia, (d) retroalimentación del profesor con toda la clase cuando comparte las observaciones ofrecidas a cada grupo y (e) reconocimiento público del trabajo bien hecho (Johnson & Johnson, 1994, 1999c).

El procesamiento de grupo con retroalimentación individual es mucho más efectivo que cuando el grupo recibe como colectivo esa retroalimentación por tres razones: (a) la motivación al logro, los logros actuales, la uniformidad de los logros en comparación con los miembros del grupo y la influencia hacia un alto logro del grupo cooperativo; (b) las relaciones positivas entre los miembros del grupo y entre los estudiantes con el profesor; y, (c) la autoestima de los alumnos y sus actitudes positivas hacia la materia de estudio. Las habilidades sociales de los miembros del equipo es una variable moderadora de la productividad de los grupos (Archer-Kath, Johnson y Johnson 1994). Además, el establecimiento de relaciones positivas entre los miembros conlleva a que los grupos trasciendan el ámbito académico porque las interacciones se fortalecen fuera de las aulas de clase.

Johnson, Johnson, Stanne y Garibaldi (1990) descubrieron que los estudiantes con niveles altos de rendimiento cuando trabajan cooperativamente en conjunto con el profesor –quien les especifica las herramientas cooperativas que deben usar– y, comparan su procesamiento grupal con el que propone el docente, alcanzan a demostrar sus capacidades como grupo y como individuos. Durante el procesamiento de grupo los miembros expresan respeto a las contribuciones de cada uno de los integrantes para que se conozca el esfuerzo que cada quien ha impreso a la tarea alcanzada porque “el procesamiento del grupo es otra de las variables moderadoras entre el aprendizaje cooperativo y el rendimiento. Las acciones individuales ayudan o no a decidir qué acciones continúan o cambian dentro del grupo” (p. 507).

Lo cierto es que no todos los grupos son “grupos cooperativos eficientes”. Cuando están presentes todos los principios o elementos esenciales, es cuando verdaderamente se habla de un grupo cooperativo efectivo como bien lo señalan Johnson y Johnson

(2014a) en donde el grupo es mas que la suma de las partes, a diferencia de los “pseudogrupos” en donde los miembros han sido asignados a trabajar cooperativamente pero no están interesados en ellos o del “grupo cooperativo tradicional” en donde los estudiantes piensan que van a ser evaluados de manera individual y no como miembros de un grupo.

En palabras de Johnston (2009), el futuro del aprendizaje cooperativo depende más del regreso a los principios básicos para una efectiva instrucción que en la innovación de una estrategia en particular. Las evaluaciones de las universidades indican que cuando los estudiantes no tienen claro lo que tienen que hacer y cómo hacerlo, el aprendizaje activo deja de ser interesante para profesores y alumnos “clear cooperative tasks requires the generation of essential content statements that specify what it is that students need to know about a topic” (p. 349).

### ***1.5. Tipos de grupos de aprendizaje cooperativo***

El aprendizaje cooperativo no sacia sus maneras en el mero hecho de las agrupaciones. Tres tipos de grupos son los que el docente puede planificar: **grupos formales**, **grupos informales** y **grupos de base**, los cuales son muy diferentes en cuanto a su implementación, la actitud del docente y, lo que es mas importante a la postre, los resultados de aprendizaje académicos y sociales.

El primero de los tipos de grupo de aprendizaje cooperativo, bautizado como **grupo formal**, se plantea desde diversas condiciones que lo determina: (a) tiene una duración que va desde una clase hasta varias semanas, (b) se puede estructurar para cualquier actividad académica y (c) procura que los alumnos se involucren activamente en la organización de materiales, su explicación, resumen e integración de las estructuras conceptuales (Johnson & Johnson, 1999c; Johnson & Johnson, 2014a; Johnson, Johnson & Holubec, 1992; Johnson, Johnson & Holubec, 1999).

Por su parte, el segundo tipo de grupos, de **aprendizaje cooperativo informal**, se caracteriza porque: (a) su duración varía entre algunos minutos y una sola sesión de clase, (b) se usa para que los alumnos se concentren en una explicación –enseñanza directa– que implique una atención especial sobre el tema a estudiar, (c) asegura el aprovechamiento cognitivo sobre el material y (d) genera un cierre o conclusión (Johnson & Johnson, 1999c; Johnson & Johnson, 2014a; Johnson, Johnson & Holubec, 1992; Johnson, Johnson & Holubec, 1999).

Por último, y no por ello menos importante, los **grupos cooperativos de base**, que tienen como propósito el apoyo mutuo de los integrantes para progresar académicamente, plantean (a) la necesidad de la heterogeneidad de sus miembros así como su estabilidad prolongada, (b) la larga duración de los mismos (todo un año) y (c) un compromiso a largo plazo entre los miembros del grupo (Johnson & Johnson, 1999c; Johnson & Johnson, 2014a; Johnson, Johnson & Holubec, 1992; Johnson, Johnson & Holubec, 1999).

Lo que sí es una constante en estos tres tipos de grupos de aprendizaje cooperativo es la labor del docente. Se trata, fundamentalmente de definir y especificar los objetivos de la actividad, en especial los objetivos sociales que le permitirán controlar los grupos de aprendizaje y diagnosticar habilidades específicas para la resolución de problemas. De esa manera, los mismos estudiantes identificarán aquellas habilidades que mejorarían el trabajo en equipo, enunciándolas como parte del proceso de enseñanza–aprendizaje y, por qué no, para analizarlas dentro de las necesidades para la implementación de una tarea (Johnson & Johnson, 1999c).

Otro reto al que se enfrenta el docente al escoger el tipo de agrupación, es el tamaño que deberá tener el grupo y, una buena manera de decidir la dimensión está en la disposición de los materiales que se requieren para la actividad.

A mayor tamaño, mayor diversidad en la gama de aptitudes, en las destrezas y en las habilidades, lo que se convierte en una mayor disposición para el procesamiento de la información y, por supuesto, una mayor diversidad de puntos de vista. Esta condición hace que si el grupo es muy numeroso, las posibilidades de alcanzar las metas que se proponga el profesor requerirán un tiempo mas prolongado, razón de más para que de acuerdo a los objetivos de aprendizaje que deba emprender el grupo y del tiempo de que se disponga, se determine su tamaño (Johnson & Johnson, 2014a).

Otro desafío del tamaño de los grupos es que mientras más pequeños sean, los miembros tendrán menos oportunidades de ocultarse con sus compañeros, por lo que la participación de todos los miembros del grupo es fundamental. Y, por supuesto, cuando mayor sea el grupo, más habilidades tendrán que mostrar cada uno de sus integrantes, puesto que “en la medida en que aumenta el tamaño del grupo, se hacen más complejas y sofisticadas las habilidades interpersonales” (Johnson & Johnson, 1999c, p. 39). Pero,

en contraparte, en tanto en cuanto el grupo se hace de un tamaño mayor, las posibilidades de interacción son menores.

Y, cuando el grupo es menor, más fácilmente se pueden detectar las dificultades que tengan los estudiantes para trabajar en conjunto.

En consecuencia, aunque los hermanos Johnson apunten que “no existe el grupo ideal” (1999c, p. 41), puesto que no versa el trabajo de los grupos por quienes son sus miembros sino por los resultados que obtienen trabajando juntos, la decisión del tamaño del grupo depende, como se ve, de los factores y condiciones que se quieran para alcanzar las metas de aprendizaje (Johnson & Johnson, 1999c).

Ahora, no se trata solamente de crear grupos de acuerdo a los resultados de aprendizaje, sino de hacer que dentro de estos grupos se generen relaciones de interdependencia. Para ello hay que asignar roles dentro de los grupos, delegar una tarea a cada uno dentro del grupo para que las funciones de liderazgo tengan sentido.

Al asignar roles se desprenden una serie de condiciones para quien asume ese encargo. Pero, como en todo proceso de enseñanza–aprendizaje, la asignación de roles requiere de un proceso paulatino. Primero, al no asignar roles, los alumnos se van acostumbrando a trabajar en conjunto. Una vez que lo saben hacer, es oportuno definir roles elementales como por ejemplo de “formación, de lector o de encargado de estimular la participación” (Johnson & Johnson, 1999c, p. 45), siempre con una rotación de esos roles para que los alumnos tengan conciencia de las diferentes funciones y, cada vez que se dé una rotación de esos roles, agregar uno nuevo, un poco más complejo, como por ejemplo, el rol de funcionamiento que versa sobre las relaciones de un aprendizaje nuevo con el anterior y de criticar ideas, no personas (Johnson, Johnson & Holubec, 1992).

Una vez que el grupo conoce las diferentes funciones de los roles, se puede dar paso a nuevos roles como los de formulación o de fermentación. El primero, comporta la obligación de permanecer en el grupo, y cambiar de idea sólo si es convencido lógicamente (las destrezas se orientan a resumir, corregir, elaborar, verificar la comprensión, explicar...).. Y, el de fermentación tiene como responsabilidad la explicación del razonamiento paso a paso y de ofrecer ideas y conclusiones propias (las destrezas se centran en criticar ideas sin criticar personas, integrar ideas diferentes en



una única conclusión, ampliar la respuesta de otro agregando información, producir varias respuestas...) (Johnson & Johnson, 1999c).

### ***1.6. Variables moderadoras que inciden en el aprendizaje cooperativo***

Johnson y Johnson (1985) son los primeros en estudiar once variables explicativas que inciden en la eficacia de los grupos de aprendizaje cooperativo en comparación con el aprendizaje competitivo e individual, agrupadas por Hertz–Lazarowitz (1985) en tres grandes categorías: variables cognitivas, variables sociales y variables instruccionales.

Lo llamativo es que esas variables moderadoras solo explican “... algo más de 20% de la varianza total manifiesta en los resultados” (Coll, 1991 citado por Gavilán y Alario, 2010, p. 167) dato que releva la necesidad de seguir indagando qué otros factores influyen en la superioridad del aprendizaje cooperativo.

#### **1.6.1 Variables cognitivas**

Al facilitar a los estudiantes situaciones de clase enriquecedoras como las discusiones en grupo, las variables cognitivas se estudian en cuanto a (a) la calidad de las estrategias de aprendizaje empleadas, (b) la controversia constructiva versus el debate y la concurrencia en la interacción y (c) el discurso oral que emerge en los grupos cooperativos.

##### *Calidad de las estrategias de aprendizaje*

La calidad de las estrategias de aprendizaje es un factor que incide en los grupos cooperativos. Las más efectivas corresponden a estrategias de búsqueda y recuperación de la información, de clasificación, formulación de ecuaciones a partir de estudios de casos y el evitar repeticiones y errores en tareas de razonamiento espacial. Cuando se comparan las estrategias que se utilizan en situaciones competitivas e individuales, los grupos cooperativos alcanzan estrategias de mayor calidad para completar sus tareas (Johnson et al 1980; Johnson & Johnson, 1981a; Skon et al 1981 citados por Johnson & Johnson, 1985) porque el proceso de discusión que ocurre en los grupos cooperativos promueve el descubrimiento y el desarrollo de estrategias cognitivas de alto nivel (Johnson & Johnson, 1985).

### *Controversia constructiva versus debate y concurrencia*

La participación en situaciones cooperativas inevitablemente produce conflicto entre los puntos de vista de los miembros de un grupo y ese conflicto que emerge en los grupos puede orientarse hacia la controversia constructiva, el debate y la concurrencia.

La *controversia constructiva* entendida como el acuerdo al que llegan los miembros de un grupo a pesar de sus diferencias iniciales, se distingue de otras estrategias de aprendizaje que se utilizan con mayor frecuencia en las aulas (la *concurrencia*, el *debate* y los *esfuerzos individuales*) no siendo la estrategia que se utilice con mayor reiteración a pesar de sus beneficios comprobados en el rendimiento y en las habilidades sociales y emocionales.

En cambio, el proceso de *concurrencia* ocurre cuando los miembros de un grupo enfatizan los acuerdos, inhiben la discusión para evitar cualquier disenso o argumento, evitando así alternativas y cursos de acción. Es una combinación entre cooperar y evitar el conflicto (Johnson & Johnson, 2009a).

En el *debate*, dos o mas personas argumentan posiciones que son incompatibles y un moderador o juez declara al ganador en función del mejor razonamiento. Como requiere refutación de otros puntos de vista, los individuos aprenden a manejar información opuesta y, esos esfuerzos individuales de estudiar los dos lados de un mismo objeto de estudio se convierten en un tópico que genera controversia pero que no indispensablemente produce un discurso oral, por lo que las conclusiones iniciales no son desafiantes en sí mismas, ya que los miembros del grupo se sienten libres de validar sus posturas iniciales (Johnson & Johnson, 2009a).

Y, por último, en los *esfuerzos individuales* la persona trabaja de manera individual sin interactuar con el otro, ya que para alcanzar sus objetivos, no requiere relacionarse ni depende de los otros (Johnson & Johnson, 1985).

En consecuencia, los estudiantes que participan en estructuras de grupo fundadas en la controversia constructiva demuestran mayor curiosidad y exploran el problema de manera mas profunda lo que induce al conflicto cognitivo (Slavin, 1989a; Tjosvold & Field, 1986) y por consiguiente, a elevar el nivel de razonamiento.

### *El discurso oral*

El discurso oral que se produce en la interacción entre los miembros del grupo es fundamental para el almacenamiento de información, como promotor de la memoria a largo plazo y el incremento del rendimiento. En los grupos cooperativos prevalecen interacciones verbales que promueven estrategias de alto nivel orientadas a ofrecer nueva información y explicaciones y a la producción de ensayos e integración de conceptos, en comparación con estrategias de bajo nivel cognitivo como es la repetición de información (Johnson & Johnson, 1983; Johnson, Johnson, DeWeerd, Lyons & Zaidman, 1983; Johnson, Johnson, Roy & Zaidman, 1984; Lyons, 1982; Roy, 1982 citados por Johnson & Johnson, 1985).

### **1.6.2. Variables sociales**

En situaciones cooperativas, las variables sociales que han llamado la atención entre los investigadores se han centrado en (a) el apoyo entre pares, regulación y retroalimentación, (b) la participación activa en el aprendizaje, (c) el apoyo psicológico y la aceptación y (d) las actitudes positivas hacia las asignaturas.

#### *Apoyo entre pares, regulación y retroalimentación*

La regulación entre pares, la retroalimentación, el apoyo y el estímulo inciden en el compromiso y la motivación de aquellos miembros del grupo que requieren una mayor orientación externa de sus progresos para completar la tarea. En las situaciones cooperativas, los miembros del grupo se benefician del esfuerzo de todos para alcanzar el rendimiento a diferencia de las situaciones individuales y competitivas. La frecuencia en el número de ayudas y tutoría es mayor en situaciones de cooperación (Armstrong, Johnson & Balow, 1981; Cooper, Johnson, Johnson & Wilderson, 1980; Johnson & Johnson, 1981a, b, 1982a, b, 1983; Johnson, Johnson, Tiffany & Zaidman, 1983 citados por Johnson & Johnson, 1985) y las expectativas de que los otros facilitan el camino y hacen un esfuerzo para alcanzar la meta, se incrementa en este tipo de aprendizaje (Johnson & Johnson, 1972; Tjosvold, Johnson & Johnson, 1981 citados por Johnson & Johnson, 1985).

#### *Participación activa en el aprendizaje*

El aprendizaje cooperativo promueve la participación mutua y los intercambios verbales porque los estudiantes tienen que compartir el material con sus compañeros. Las

evidencias indican, a favor de las situaciones cooperativas, cuando se comparan con competitivas e individuales, que los alumnos tienden a expresar sus ideas en grupos pequeños de trabajo o frente a toda la clase, lo que favorece el deseo de compartir las propias ideas y se traduce en una mayor disposición a ofrecer soluciones (Johnson & Johnson, 1983; Johnson, Johnson, Roy & Zaidman, 1984; Johnson, Johnson, Tiffany & Zaidman, 1983 ver todos estos estudios en Johnson & Johnson, 1985) y en sentimientos positivos hacia la experiencia instruccional (Garibaldi, 1979; Gunderson & Johnson, 1980; Johnson, Johnson, Johnson & Anderson, 1976; Wheeler & Ryan, 1973 ver todos estos estudios en Johnson & Johnson, 1985).

#### *Apoyo psicológico y aceptación*

A diferencia de las estructuras competitivas e individuales, los estudiantes cuando trabajan en grupo perciben que los otros compañeros se preocupan por el aprendizaje, se ofrece ayuda, se escucha con atención y se es aceptado “tal cual como eres” (Cooper et al, 1980; Gunderson & Johnson, 1980; Johnson, 1983; Johnson & Ahlgren, 1976; Johnson, Bjorkland & Krotee, 1984; Johnson et al, 1980; Johnson & Johnson, 1981a, b, 1982b, 1983, 1984a, b, c; Johnson, Johnson & Anderson, 1978; Johnson, Johnson, Johnson & Anderson, 1976; Johnson, Johnson, Tiffany & Zaidman, 1983; Johnson, Johnson, Roy & Zaidman, 1984; Skon et al 1981; Tjosvold, Marino & Johnson, 1977, ver todos estos estudios en Johnson & Johnson, 1985).

#### *Actitudes positivas hacia las asignaturas*

Las actitudes y experiencias hacia las asignaturas, contenidos y temas que se trabajan en situaciones de aprendizaje cooperativas son mucho más positivas (Garibaldi, 1979; Gunderson & Johnson, 1980; Johnson & Johnson, 1979; Johnson, Johnson & Skon, 1979; Lowry & Johnson, 1981; Smith et al, 1981; Wheeler & Ryan, 1973 ver en Johnson & Johnson, 1985) así como aumentan la satisfacción con la clase mejorando el clima del aula (Johnson y Johnson, 1999c).

### **1.6.3. Variables instruccionales**

Finalmente, las variables instruccionales también han sido foco de interés para el estudio de su incidencia en la cooperación como por ejemplo (a) el tamaño de los grupos, (b) la formación de los grupos, (c) el tipo de tarea, (d) el tiempo de resolución de la tarea, (e) los niveles de habilidades de los miembros del grupo, (f) el

reconocimiento y asignación justa de la tarea, (g) la evaluación del aprendizaje y la revisión del funcionamiento de grupos y (h) los roles ‘docente’ y ‘alumno’ en el aprendizaje cooperativo.

#### *El tamaño de los grupos*

Los grupos para las actividades de aprendizaje cooperativo dependerán precisamente de eso, del tipo de actividad, puesto que si los grupos son numerosos, la responsabilidad se puede diluir o, en el peor de los casos, ocultar frente a quienes pueden tener más control de los contenidos. Dependerá del número de alumnos que se tenga en una clase, pero entre dos y cinco alumnos es un buen tamaño para iniciar las actividades de aprendizaje cooperativo (Gavilán y Alario, 2010).

#### *La formación de los grupos*

Una vez que se tiene definido el tamaño de los grupos de acuerdo al tipo de actividad que se va a desarrollar, los grupos pueden ser heterogéneos u homogéneos pero se sugiere la conformación de grupos con niveles distintos de conocimientos, es decir, que los alumnos tengan habilidades diversas, que representen a los diferentes niveles de la clase (Johnson y Johnson, 2014b). Esta asignación puede realizarse: (a) mediante muestreo aleatorio estratificado, (b) preguntando a los alumnos con quienes les gustaría trabajar para luego de acuerdo a los criterios del profesor agrupar a los estudiantes, (c) haciendo pasar a los alumnos por diferentes pruebas de actitud y aptitud. Lo que si se debe considerar es usar varios criterios para conformar los grupos y no dejar que siempre sean los propios alumnos los que elijan con quienes trabajar o no (Gavilán y Alario, 2010).

#### *Tipo de tarea*

El tipo de tarea es una variable que ha sido estudiada con atención por considerarse como un factor explicativo (Garibaldi, 1979; Johnson & Johnson, 1979; Johnson, Johnson & Skon, 1979; Johnson, Skon & Johnson, 1980; Skon, Johnson & Johnson, 1981 ver en Johnson & Johnson, 1985). Sin embargo, no hay un tipo de tarea específica que favorezca el aprendizaje cooperativo sobre el competitivo e individual, pero sí son especiales aquellas que están orientadas al logro, la resolución de problemas verbales y espaciales, la retención, la memoria, las habilidades motoras y, en donde el alumno debe activar sus conocimientos previos para luego valorar la situación y finalmente predecir lo que ocurrirá.

### *Tiempo de resolución de la tarea*

No solo el tiempo de resolución de la tarea es mayor en situaciones cooperativas cuando se compara con situaciones competitivas e individuales, sino que además, el esfuerzo también se incrementa cuando “los integrantes de los grupos cooperativos parecen pasar mas tiempo ocupados en las tareas que los demás estudiantes (Johnson y Johnson, 1999c, p. 287). Los alumnos internalizan las conversaciones de clase a través de una reflexión metacognitiva porque “we can think because we can talk, and we think in ways we have learned to talk” (Bruffee, 1984, p. 640).

### *Niveles de habilidades de los miembros del grupo*

La composición de los grupos según los distintos niveles de habilidades de cada uno de los miembros del equipo es otra condición que influye en los resultados académicos. Diversas investigaciones han estudiado el efecto positivo de la combinación de grupos homogéneos o heterogéneos de estudiantes en contraste con situaciones competitivas e individuales (Armstrong et al, 1981; Johnson, Johnson, Roy & Zaidman, 1984; Martino & Johnson, 1979; Nevin et al, 1982; Smith et al 1981, 1982, 1984; Skon et al 1981 ver todos los estudios en Johnson & Johnson, 1985), pero no hay claridad en cuanto a si los estudiantes con niveles bajos y medios de rendimiento, se benefician del trabajo grupal con distintos niveles de desempeño. En cambio, sí hay evidencias que demuestran que estudiantes con un alto rendimiento no se benefician del trabajo con estudiantes de medio y bajo nivel de desempeño (Johnson & Johnson, 1985).

### *Reconocimiento y asignación justa de la tarea*

Cuando los estudiantes trabajan cooperativamente, el reconocimiento es justo porque se valora el esfuerzo individual del estudiante y su implicación en el grupo. De esta forma, los alumnos perciben que si todos se esfuerzan tienen la misma posibilidad de alcanzar el éxito.

El sistema de recompensas es equitativo porque reconoce el esfuerzo individual y grupal. Los estudiantes prefieren este sistema cooperativo para la evaluación de la tarea en comparación con un sistema que favorezca el aprendizaje competitivo o individual (Johnson, 1976; Johnson & Johnson, 1976; Johnson, Johnson & Bryant, 1973; Johnson, Johnson, Johnson & Anderson, 1976; Johnson, Ryan & Schroeder, 1974 ver todos estos estudios en Johnson & Johnson, 1985). Los alumnos en situaciones competitivas al

inicio de la actividad, perciben el sistema de reconocimientos y recompensas como justo, pero no así una vez que finaliza (Deutsch, 1979; Johnson & Johnson, 1975b ver en Johnson & Johnson, 1985).

#### *La evaluación del aprendizaje y la revisión del funcionamiento de los grupos*

La evaluación debe ser notificada al alumno antes de emprender la tarea de evaluar. El estudiante debe conocer cómo será valorado y bajo qué criterios.

Para Gavilán y Alario (2010) dentro de los criterios que se pueden evaluar están también las habilidades sociales, tan importantes dentro de las bases filosóficas del aprendizaje cooperativo puesto que, además de los alcances académicos que obtengan los alumnos, los valores de solidaridad y elementos de carácter personal como la autoestima y la mejora de la salud psíquica forman parte de este enfoque. Se comprende al estudiante como un ser global, en formación con otros.

**Evaluar el rendimiento académico** comporta entonces, una “serie de reflexiones previas a la evaluación académica del alumnado” (Gavilán y Alario, 2010, p. 146) que pasan por la especificidad de los objetivos que se pretenden alcanzar así como el diseño de la tarea para lograrlos, mostrar y aclarar los baremos de evaluación, hacer la labor de comunicación al alumno de esos objetivos y criterios y, por supuesto, emprender el seguimiento adecuado a estos grupos estudiantiles cooperativos.

Evidentemente, para hacer ese seguimiento de la actividad del alumnado, existe un abanico de posibilidades explicada muy bien por estos autores. Por ejemplo, la observación atenta por parte del profesor de la actividad de los estudiantes, las formas como se interrelacionan y la calidad de las explicaciones que se ofrecen entre ellos. Con este tipo de observación se pueden tomar apuntes de cómo se desarrolla el aprendizaje individual con respecto al colectivo y cómo es la participación de cada uno de los miembros del grupo, con la finalidad de corregir los errores que pudieran estarse gestando en el seno de un grupo (Gavilán y Alario, 2010).

Otra forma es preguntar en clase. Al hacerlo, el profesor se percató de los niveles de comprensión y conocimientos que se adquieren con la tarea. Si se hace con todos los miembros de un grupo, la recompensa se materializa con puntuaciones para todos y cada uno de los integrantes.

También se puede hacer la revisión conjunta de las actividades trabajadas en casa (trabajo autónomo) que, aunque consume mucho tiempo de la clase, permite conocer

cómo es la implicación individual de las tareas asignadas, el tiempo que dedican a ellas, saber si la comprensión es suficientemente válida, verificar que todos los miembros del grupo han realizado sus deberes y, es más, que hayan entendido la asignación y, al hacerlo, puntuar favorablemente el proceso y el producto entregado.

Una vez que se ha pasado este proceso colectivo, los alumnos deben demostrar individualmente que dominan los contenidos y, para ello, se realiza una prueba en la que, además de demostrar el conocimiento individual, se coloca como estímulo las calificaciones que como grupo se obtienen de modo colectivo. De esta manera, desde el punto de vista personal, se comprueban los resultados entre todos los miembros del grupo porque se autoevalúan, es decir, el examen se realiza dos veces: la primera individualmente y, la segunda en grupo. Como se puede prever, la autoevaluación, como la coevaluación son parte importante del aprendizaje cooperativo.

Finalmente, dentro del aprendizaje cooperativo, hay que **calificar** y para hacerlo, hay varias consideraciones según Gavilán y Alario (2010).

La primera, es que el sistema de puntuación debe ser equilibrado para que **todos los miembros de un grupo reciban la misma puntuación**, es decir, la que sea producto de la media de todos los integrantes, lo que deja en evidencia quienes aportaron más al proceso de aprendizaje, así como quienes restaron puntuación al grupo.

Otra forma, la segunda, es sumar las puntuaciones y ponderarlas en una escala de 10 puntos. De esa forma no se pone a la vista de todos las falencias del grupo, o las muchas diferencias que podrían existir. Una tercera vía es que el resumen de la actividad es el que se califica y esa nota, al ser producto del trabajo en conjunto es la que compartirá el equipo.

La cuarta manera es que cada miembro desarrolla una parte del trabajo, pero esta debe ser evaluada y aprobada por cada uno de los integrantes del equipo, dándole peso a cada uno y siendo responsables todos de la calificación.

La quinta forma es puntuar sobre la calificación más baja del grupo, lo que pretende que todos los miembros se vuelquen a desarrollar la actividad de tal forma que ninguno tenga un bajo rendimiento.

La otra gran opción es que los **miembros del grupo reciban puntuaciones diferentes** lo que también comporta variantes. Una es que los estudiantes conserven su puntuación y se les sume la media del grupo, lo que implica un necesario apoyo para que la nota del



grupo sea la más alta posible. La otra es que cada miembro conserve sus puntos más algunos que se hayan acordado como mínimos si se realiza la tarea. Lo importante es que cualquiera de las metodologías de puntuación que se adopten, sean discutidas previamente entre el profesor y los alumnos antes de emprender la tarea (Gavilán y Alario, 2010).

### *Los roles 'docente' y 'alumno' en el aprendizaje cooperativo*

En el aprendizaje cooperativo la asignación de roles, inclusive el asumir los roles, tiene enfoques diversos de acuerdo a los autores que lo tratan. En las aulas tradicionales, plantea Kagan (1985), los estudiantes por la propia configuración del aula, adoptan el rol de estudiantes. Pero, en las actividades de aprendizaje cooperativo, el alumno puede asumir roles diversos y cambiar inclusive sus autoconceptos al convertirse en tutores, consultores, investigadores y presentadores.

En el caso de los docentes, al tener estudiantes que asumen el rol de profesores y maestros, la situación de su posición tradicional cambia (Kagan, 1985).

Así, por ejemplo, Slavin con los métodos que se agrupan bajo el paraguas del *Student Team Learning* (STL), postula que es el docente el que elige el material que se usará, define los grupos y, por supuesto, realiza la labor de evaluación de esos grupos (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

Por su parte, los hermanos Johnson, cuyo método fundamental es *Learning Together* (LT), coinciden con Slavin en las tareas que el docente debe asumir, pero son los propios materiales la búsqueda de la solución de los problemas, lo que requiere del consenso del grupo para solventarlos.

Por último el *Group Investigación* (GI), desarrollado por Sharan, plantea que es el alumno quien trabaja y toma decisiones en proyectos abiertos, lo que hace que el rol del docente sea más bien de un facilitador de la tarea, antes que un director de la misma, lo que propone una relación mucho más democrática y exige de mayor autonomía de los alumnos (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

La tabla 2 es el resumen de esa discusión comparativa entre los tres autores mencionados

Tabla 2: Rol docente y del alumno en los métodos cooperativos

<b>Factores</b>	<b>Slavin</b>	<b>Johnson y Johnson</b>	<b>Sharan</b>
<b>Rol docente</b>	Determina los objetivos y medios para aprender	Determina los objetivos y medios para aprender	Asiste a los alumnos en el proceso de decisión de los objetivos y medios de aprendizaje
<b>Rol alumno</b>	Aprende el material que le es asignado para competir con otros grupos	Trabaja sobre los problemas y se responsabiliza del aprendizaje propio	Planea, ejecuta y evalúa el proyecto de aprendizaje con el docente
<b>Contenidos currículo</b>	Material para ser aprendido sobre todo mediante la repetición y la memoria	Problemas que pueden propiciar el aprendizaje y la toma de decisiones	Temas comprensivos que integran el desarrollo de contenidos académicos

Fuente: León del Barco, Gozalo, Felipe, Gómez y Latas (2005, p. 84)

Como cierre, luego de la revisión de las variables principales que inciden en la eficacia del aprendizaje cooperativo y sin considerarse un listado acabado de factores explicativos, las habilidades sociales y de interacción son las fundamentales en la eficacia de esta metodología (Echeita, 1995; Johnson, Johnson, Stanne y Garibaldi, 1990) porque

En el aprendizaje cooperativo no basta con dejar que los alumnos se pongan a trabajar en grupo o con promover la interacción entre ellos para obtener, de manera inmediata, unos efectos favorables sobre el desarrollo, la socialización y el aprendizaje. Lo importante no es la cantidad de interacción sino la calidad de la misma. Las personas que se unen para trabajar adquieren todas las características de los grupos sociales: presión del grupo, conflictos, adopción de roles... Qué ocurre si los miembros del grupo no se conocen, si no hay un clima de confianza entre ellos, si no saben cooperar, si no son responsables, si no se comunican, si hay marginación, chivos expiatorios, actitudes de conformismo, disputas que mantienen a grupos antagónicos en conflicto, pasividad extrema, etc. La toma de conciencia de estos problemas, plantea la búsqueda de soluciones que den lugar a interacciones entre los alumnos ricas y constructivas en cuanto a sus efectos (León del Barco, 2006, p. 106).

La productividad y el nivel de aprendizaje del grupo estará mediado por la madurez del conocimiento que exige la dinámica social: un grupo que ha experimentado situaciones diversas en contextos y tareas diferentes logrará un mejor desempeño. La cooperación no se aprende por intuición, requiere reflexión, sistematización, seguimiento y una continua evaluación.

### ***1.7. Métodos cooperativos***

El tema de la clasificación de los métodos cooperativos debe pasar indispensablemente por tener un mapa de términos. Debido al origen americano de esta metodología, la

mayoría de las investigaciones, disertaciones doctorales y textos de referencias se han difundido en inglés, lo que obliga su traducción a otros idiomas. Este ejercicio –siempre complejo– trae como consecuencia que la literatura, sobre todo en materia de siglas, presente diferentes acepciones.

Para lograr comprender estas disparidades de nomenclatura, se ha elaborado un glosario de términos en Inglés – Español, que a los efectos permite clasificar a los métodos del aprendizaje cooperativo que provienen de la literatura revisada en las dos lenguas vehiculares de las que se ha nutrido esta tesis (ver anexo 1). Sin embargo, en el cuerpo del trabajo se hará referencia en el idioma inglés a los distintos métodos cooperativos por entender que son cómo se conocen e identifican en la literatura científica.

Autores referentes como Slavin y Kagan y Kagan coinciden en la eficacia del aprendizaje cooperativo y de las bondades de este tipo de enseñanza pero difieren en cuanto al rol del docente

Slavin emphasizes team learning and provides a sequential pattern to be followed by the educator. The patterns are adaptable, so they can be modified for given applications (...) Kagan maintains that an educator needs to have a group of structures in which any learning can be placed. Kagan is similar to Slavin in this approach except that Kagan is trying to prevent the innovation from being bound to a particular set of information so that the teacher is able to make choices about how to present information (Randolph, 1992, pp. 38-39).

Y en lo que refiere a cómo agrupar los métodos cooperativos no es precisamente un ámbito en el que los autores armonicen. En el año 1995, Slavin sugiere la siguiente clasificación que se presenta en la tabla 3:

Tabla 3: Clasificación de los métodos cooperativos según Slavin

<b>STUDENT TEAM LEARNING (STL)</b>
Student Teams–Achievement Divisions (STAD)
Teams–Games–Tournament (TGT)
Team Assited Individualization (TAI)
Cooperative Integrated Reading and Composition (CIRC)
<b>GROUP INVESTIGACIÓN (GI)</b>
<b>COMPLEX INSTRUCTION (CI)</b>
<b>LEARNING TOGETHER (LT)</b>
<b>JIGSAW</b>
<b>JIGSAW II</b>
<b>DYADIC METHODS</b>

Fuente: Slavin (1995, p. 12)

Mientras que sus colegas Kagan y Kagan, en el año 2009 sugieren esta otra, en donde se observa en la tabla 4 que incluye al igual que Slavin métodos tradicionales pero a algunos de ellos los clasifica de una manera diferente:

Tabla 4: Clasificación de los métodos cooperativos según Kagan

<b>JIGSAW DESIGNS</b>	
Original Jigsaw	
Jigsaw II	
Variación de Jigsaw según Kagan	<ul style="list-style-type: none"> <li>▪ Within–Team Jigsaw</li> <li>▪ Pairs</li> <li>▪ Team Jigsaw</li> <li>▪ Partner Expert Group Jigsaw</li> <li>▪ Double Expert Group Jigsaw</li> <li>▪ Workstation Jigsaw</li> <li>▪ Leapfrog Jigsaw</li> </ul>
Controversy Jigsaw	
Jigsaw Problem Solving	
Partners	
<b>COOPERATIVE INVESTIGATION</b>	
Group Investigation (GI)	
Co–op Co–op	
Co–op Jigsaw	
Complex Instruction (CI)	
<b>MASTERY DESIGNS</b>	
Color–Code	
Co–op Cards	
Student Teams–Achievement Divisions (STAD)	
Teams–Games–Tournament (TGT)	
Team Assisted Individualization (TAI)	
Cooperative Integrated Reading and Composition (CIRC)	
Co–op Centers	
<b>LEARNING TOGETHER (LT)</b>	

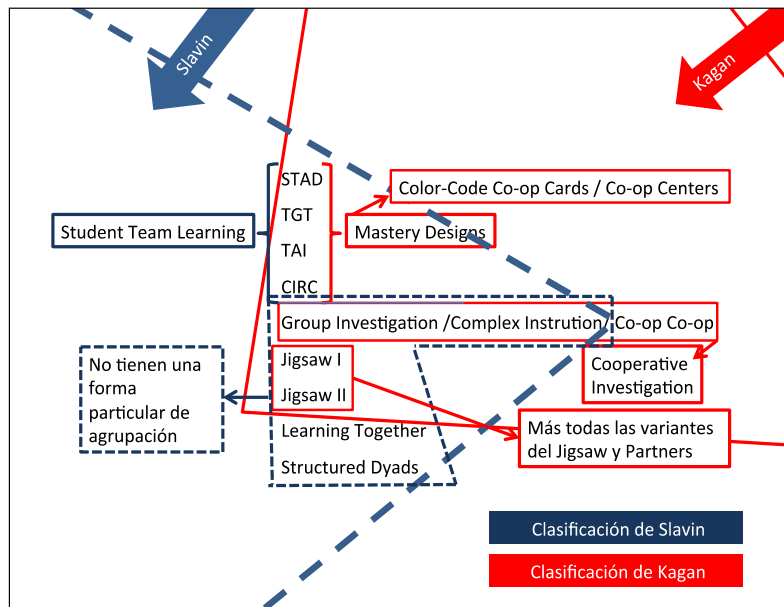
Fuente: Kagan & Kagan (2009, p. 17.2)

Al comparar ambas propuestas, Slavin, cataloga los métodos de acuerdo a su propia naturaleza y engloba a cuatro de los métodos –STAD, TGT, TAI y CIRC– bajo el nombre de *Student Team Learning*. Los demás métodos no tienen para este autor una forma particular de clasificarse– GI, CI, LT, *Jigsaw*, *Jigsaw II* y *Dyadic Methods*, al contrario de lo que proponen Kagan y Kagan, que ordenan a los métodos *Student Team Learning* –STAD, TGT, TAI y CIRC– como *Mastery Designs* y asigna a los demás métodos como GI y CI bajo el nombre de *Cooperative Investigation*.

Es interesante notar que Slavin coloca en su clasificación general el trabajo por parejas *Structured Dyads*, mientras que Kagan y Kagan lo incluye como una variante del *Jigsaw Designs*, en donde el *Jigsaw* y *Jigsaw II* pertenecen a esta misma categoría e incorpora nuevos métodos como el Co–op Co–op y todas sus variantes bajo la categoría

*Cooperative Investigation.* La ilustración 1, muestra gráficamente las fronteras que han delimitado los dos autores:

Ilustración 1: Clasificación de los métodos cooperativos según Slavin y Kagan



Fuente: elaboración propia

Ahora bien, de los métodos que se han mencionado ¿cuáles han sido creados para la universidad? La tabla de seguidas (5) da cuenta de los métodos utilizados con mayor frecuencia en el ámbito universitario, sus creadores y los años o década en que se han diseñado:

Tabla 5: Principales métodos cooperativos en la universidad

Método	Autor(es)	Años/Década
Learning Together (LT)	Johnson y Johnson	Mediados de los 60
Jigsaw	Aronson y Cols.	Finales de los 70
Jigsaw II	Slavin	1986
Group Investigation (GI)	Thelen, Sharan y Sharan	Mediados de los 70
Co-op Co-op	Kagan y Kagan	Mediados de los 80
Peer Collaboration	Damon y Phelps	1989
Peer Tutoring	Topping y colaboradores	1996
Reciprocal Peer Tutoring (RPT)	Fantuzzo y colaboradores	Década de los 90
Scripted Cooperation (SC)	O'Donnell, Dansereau y Lambiotte	Finales de los 80
Peer Teaching	Goldschmid y Goldschmid	1976
Teams-Games-Tournament (TGT)	DeVries & Edwards	Inicios de los 70
Student Teams-Achievement Divisions (STAD)	Slavin	Finales de los 70
Informal Methods	Kagan y Kagan	1992
Cooperative Learning and Computers	Varios autores	Década de los 90

Fuente: adaptación de Johnson & Johnson (2002b, p. 6); Kagan & Kagan (2009, p. 17.2) y Slavin (1995, p. 12).

Como se observa, el desarrollo e investigación de estas formas de aprendizaje cooperativo se han extendido a partir de la década de los 70. Hay métodos que se pueden usar para varios propósitos, de hecho, una de las fortalezas del aprendizaje cooperativo es su adaptabilidad para los distintos entornos de aprendizaje. Pero, como todas las disciplinas tienen su especificidad, no siempre se puede educar de la misma manera y es por esa razón, que se crean métodos mas generales y otros específicos.

La versatilidad de lo genérico en su aplicabilidad "... en una amplia diversidad de materias, grados, aulas y escuelas" (Slavin, 1999a, p. 123) habla de la fortaleza de los métodos, de su difusión y popularidad. Por otra parte, pensar que los distintos métodos de aprendizaje cooperativo, cada uno con sus matices "...resultan adecuados para *todas* las materias y grados escolares (...) para cualquier materia o grado en *particular*" (Slavin, 1999a, p. 123) es una ilusión.

Las próximas páginas son el reflejo solo de la variedad de métodos cooperativos que se aplican con mayor regularidad en el ámbito universitario, del esfuerzo por mejorar la técnica, en fin, una muestra de la especificidad de la puesta en marcha en esta área de conocimiento.

### 1.7.1. *Student Team Learning* (STL)

Bajo la denominación de *Student Team Learning* pertenecen un grupo de métodos cooperativos desarrollados por la Universidad Johns Hopkins que comparten en común

la utilización de metas grupales en donde el éxito solo se logra en la medida en que cada uno de los miembros del equipo domina la información. De estos métodos, el TAI y el CIRC se han desarrollado exclusivamente para la etapa de Educación Primaria. El primero centrado en matemáticas y el segundo, en lectura y escritura. Otros como TGT, STAD y *Jigsaw II* pueden aplicarse desde primaria hasta la universidad, siendo estos dos últimos los que el mayor número de investigaciones reportan.

#### *1.7.1.1. Student Teams–Achievement Divisions (STAD)*

El STAD es un método creado en los años 70 por Slavin, el cual se ha aplicado en distintas edades, desde los 7 hasta los 17 años e inclusive con estudiantes universitarios, en asignaturas muy variadas como Matemáticas, Lengua, Ciencias y Sociales. Tiene características que permiten recompensas de acuerdo al rendimiento y a las mejoras que se puedan alcanzar porque las tareas están pensadas para que el estudiante aprenda dentro del grupo, donde el trabajo del equipo no está acabado hasta que todos los miembros hayan adquirido el dominio del material que va a ser estudiado (Slavin, 1999b). El número de integrantes por grupo lo constituyen entre 4 y 5 alumnos y son una muestra representativa del aula de clase, lo que quiere decir que los grupos son heterogéneos.

Existen, en palabras del propio Slavin (1999b) tres grandes conceptos fundamentales en el STAD: las recompensas grupales, la responsabilidad individual y la igualdad de oportunidades para alcanzar el éxito. Otro componente esencial es que los grupos no están compitiendo entre sí porque solo deben alcanzar la meta al finalizar la tarea. Así, la responsabilidad individual se referirá al compromiso que adopte cada uno de los miembros con el resto del equipo para obtener el éxito. El foco de la actividad de los miembros del equipo debe ser la tutoría entre unos y otros para asegurarse de que todos los integrantes están suficientemente preparados para superar la prueba. De esa manera, la igualdad de oportunidades para alcanzar el éxito significa que todos los estudiantes contribuyen con su grupo para obtener ese logro.

En cuanto a la evaluación, se parte de una “situación base” que el estudiante tendrá que superar si quiere sumar puntos para su equipo. Así se garantizan dos condiciones fundamentales: esfuerzo y mejora del desempeño. El esfuerzo se traduce en la ‘superación consigo mismo’. Por ello “cualquier alumno puede aportar el máximo de

puntos a su equipo en este sistema pero ninguno puede hacerlo sin esforzarse” (Slavin, 1999a, p. 86).

#### *1.7.1.2. Teams–Games–Tournament (TGT)*

El TGT fue desarrollado por DeVries y Slavin en 1978 aunque sus inicios datan de años anteriores (DeVries, 1976; DeVries, Edwards & Slavin, 1978; DeVries, Edwards & Wells, 1974). No difiere en esencia del STAD porque comparten la misma dinámica general, es decir, la preparación de la clase, la conformación de los equipos y el reconocimiento por equipos, no así los juegos y torneos (Knight & Morton, 1990). Lo que no es común es la condición de “superación individual” que sí se observa en el STAD. En este método, su nombre nos indica que los estudiantes participan en torneos académicos en donde los representantes de cada grupo compiten con otros alumnos de otros equipos con niveles similares de desempeño. Como la diferencia principal radica en este aspecto, algunos profesores han optado por combinar ambos métodos. De esta forma aplican la estructura del STAD pero con torneos eventuales (Slavin, 1980).

Una característica distintiva de este método es que los estudiantes, como representantes de sus equipos, compiten con estudiantes de otros grupos cuyos resultados académicos han sido similares a los suyos (Hagelgans, et al. 1995). Como método idéntico al STAD difiere en que las pruebas son reemplazadas con torneos a manera de juegos y los resultados individuales se contabilizan bajo el criterio de sistema de puntaje o *bumping system*.

El *bumping system* pretende igualar las oportunidades de obtener puntos para el grupo al que se pertenece, ya que las tablas de torneo son homogéneas al considerar las habilidades del nivel de los integrantes. Inicialmente el profesor ha asignado a los estudiantes las hojas de trabajo para el torneo. Luego, el *bumping system* reasigna a los estudiantes para las mesas del torneo asumiendo que en cada torneo el mayor rendimiento hace que se avance o se ascienda en la escala y, a menor rendimiento se descende en esa escala (Kagan, 1985).

El TFG cumple con cuatro grandes principios: (a) el juego se incorpora en la estructura normal de recompensas de la clase; (b) el juego refuerza la experiencia de aprendizaje no reemplazando al proceso de adquisición de información; (c) la competencia individual incide en la competencia grupal y, (d) el éxito de los juegos se asume como una responsabilidad que busca adquirir herramientas de conocimiento (DeVries, 1976).



### 1.7.1.3. *Jigsaw II*

El método de aprendizaje cooperativo que sienta sus bases en la suma de las partes, se escribe en inglés *Jigsaw* y ha sido creado por Aronson et al. en 1978 bajo el nombre de Rompecabezas original, el cual ha sufrido modificaciones en el tiempo como *Jigsaw II*, *Jigsaw III* y *Jigsaw IV*, entre otros. De todos, el *Jigsaw II* es el que cumple con los criterios del *Student Team Learning* porque a diferencia de las otras variantes y en especial, del *Jigsaw I*, los estudiantes reciben toda la información del tema y el rendimiento individual contribuye al desempeño del grupo al utilizar incentivos para la cooperación (Aronson & Patnoe, 2011).

En el *Jigsaw II*, los alumnos trabajan en equipos heterogéneos al igual que en los métodos STAD y TGT. A cada alumno se le asigna capítulos o unidades que debe leer. Esta lectura se acompaña con “hojas expertas” que contienen diferentes tópicos relacionados con la lectura asignada. Una vez que han terminado la lectura, los alumnos de diferentes grupos que tienen el mismo tópico se reúnen en un “grupo experto” en donde se discute el tema estudiado. Luego, regresan a sus grupos originales para comentar lo que han hablado en el grupo experto. Los estudiantes tienen que presentar evaluaciones que se relacionan con todos los tópicos trabajados en el equipo (Slavin, 1999a).

Al igual que en el STAD, los puntajes que aportan al grupo provienen de la superación individual. Los equipos que logren un mayor rendimiento obtendrán recompensas u otras maneras de reconocimiento. La interdependencia de los miembros del grupo está presente: todos dependen de todos para lograr evaluaciones satisfactorias (Slavin, 1999a).

La tabla resumen 6 compara muy bien las diferencias y semejanzas entre los métodos mas frecuentes que se agrupan bajo la categoría *Student Team Learning*:

Tabla 6: Cuadro comparativo *Jigsaw II*, STAD y TGT

	JIGSAW II	STAD	TGT
Número de alumnos	Cuatro o cinco alumnos		
Conformación del grupo	Heterogéneos		
Estructura meta	Alta: los participantes solo pueden conseguir su objetivo si los demás también lo consiguen		
Estructura tarea	Alta	Baja	Baja
Estructura recompensa	Alta: existen recompensas e incentivos a la cooperación basadas en los aprendizajes individuales. Existe competencia intergrupal para conseguir dichas recompensas		
Interdependencia alumnos	Se consigue una gran dependencia entre los alumnos debido sobre todo a la estructura de tarea, a la estructura de meta y a la estructura de recompensa		
Responsabilidad alumnos	Alta: motivada por la estructura de tarea, de meta y de recompensa. Los alumnos reciben información sobre su contribución a las puntuaciones grupales		
Tipo de evaluación	Pruebas individuales sobre todo el tema		Juegos competitivos, torneos formados por mesas de tres alumnos de rendimiento similar pero de diferentes grupos
Contribuir puntuación global	Existe igual oportunidad de contribuir a la puntuación grupal, pues ésta se hace en función del progreso del alumno, independientemente de su rendimiento		

Fuente: adaptación de León del Barco, Gozalo, Felipe, Gómez y Latas (2005)

### 1.7.2. Cooperative Investigation

Existen métodos especializados en contenido y conocimientos adquiridos a través de la investigación como GI, Co-op Co-op y *Complex Instruction* (CI) siendo los dos primeros pensados para la enseñanza universitaria y con una trayectoria mayor en investigación el primero de ambos. A estos, se llaman *Cooperative Investigation* o métodos de investigación orientados al aprendizaje de conocimientos y destrezas en donde la evaluación grupal proviene de un proyecto grupal.

No solo valoran las destrezas básicas y el conocimiento sino cómo los estudiantes se desarrollan como personas. La fuerza se centra en qué es estudiar y cómo se aprende. El estudiante aprende porque quiere comprender (Kagan, 1985). Los alumnos aprenden a ser alumnos porque tienen que formular preguntas, buscar fuentes de información y solucionar problemas usando el análisis y la síntesis. Son mucho más exigentes para estudiantes con niveles de rendimiento bajo debido a la necesidad de la utilización de estrategias de alto nivel cognitivo y eso requiere del uso de la expresión creativa y la verbalización de las opiniones de cada miembro del grupo (Kagan, 1985).

#### 1.7.2.1. *Group Investigation (GI)*

Slavin (1999a) reconoce que es el método de especialización que ha sido más investigado y exitoso. Las bases conceptuales descansan en la pedagogía de John Dewey pero en los últimos años el equipo de investigación de Israel, conformado por Shlomo Sharan, Yael Sharan y Rachel Hertz–Lazarowitz, son quienes llevan el liderazgo en el dominio de la técnica. Se trata de un método muy versátil que puede aplicarse en diferentes edades y en una variedad de situaciones.

Cuando los estudiantes se unen a estudiar en pequeños grupos, se pueden ayudar unos a otros y, al mismo tiempo, desarrollar una disciplina de autodirección y responsabilidad en sus aprendizajes (Sharan & Sharan, 1999). Es en ese sentido que los docentes, al incorporar a su repertorio de métodos de enseñanza los procedimientos del trabajo cooperativo, encuentran que los GI son una plataforma en la que los estudiantes utilizan todas las herramientas adquiridas en los métodos de aprendizaje cooperativo, los amplían (Sharan & Sharan, 1990) y son una manera, a la vez, de demostrar sus habilidades interpersonales. Les permite oponer al grupo sus hallazgos, la forma en cómo planifican la investigación, juntos, con la ayuda del profesor, colaborando y evaluando sus esfuerzos interpersonales (Joyce & Weil, 1986; Sharan & Hertz–Lazarowitz, 1980; Sharan & Sharan, 1992; Thelen, 1981 citados por Sharan & Sharan, 1999; Quinn, 2006).

Cuatro íes caracterizan al *Group Investigation*: investigación, interacción, interpretación e intrínseca motivación (Sharan & Sharan, 1999; Quinn, 2006). Las etapas para la puesta en práctica de este método se resumen en: (a) la clase determina los subtópicos y se organiza en grupos de investigación, (b) los grupos planifican sus investigaciones, (c) los grupos desarrollan sus proyectos de investigación, (d) los grupos planifican sus presentaciones, (e) los grupos presentan a toda la clase y por último, estudiantes y profesores evalúan los proyectos.

#### 1.7.2.2. *Jigsaw I*

¿Por qué el *Jigsaw II* pertenece a la categoría de métodos bajo la denominación *Student Team Learning* y no así el *Jigsaw I*?

El *Jigsaw I* creado por Aronson et al. en 1978, también llamado rompecabezas original, consiste en dividir un tema largo en subtemas sin utilizar material ya existente, los miembros del grupo no tienen acceso a toda la información y no aplica el sistema de

puntuación del STAD. De esta forma la técnica exige que los miembros del grupo comprendan que el otro es una fuente de información, en donde la competencia individual es incompatible con el éxito, el lograr la meta se alcanza a través de conductas cooperativas y en donde cada miembro aporta “a piece of vital information that was not readily available except from that individual student” (Aronson & Patnoe, 2011, p. 8).

Los estudiantes, con los materiales del tema subdividido se reúnen en un grupo heterogéneo conformado entre 3 y 6 miembros. A cada uno de los alumnos se le encarga la lectura de una “hoja de experto” y, como cada hoja pertenece a un tópico, se reunirán los expertos en grupos por tópicos. Reunidos los expertos se discute la “hoja de expertos”, se busca información que pueda solventar las inquietudes y se vuelve al grupo original, donde cada uno de los miembros es experto en una de las áreas compartiendo así el conocimiento adquirido (Clarke, 1999).

Uno de los requisitos de este método es que solo puede utilizarse si lo que se pretende estudiar está escrito en forma de narrativa. La mayor dificultad se encuentra en que el material debe estar diseñado para que pueda estar dividido en partes independientes. Esto implica que para la comprensión total del tema, los alumnos dependen de lo que sus compañeros le enseñen.

Es apropiado para asignaturas como Sociales, Literatura, algunos aspectos de Ciencias y otras áreas de conocimiento en donde prevalece el aprendizaje del conocimiento más que la habilidad (Doymuş, Karacop & Simsek, 2010). La evaluación se lleva a cabo sobre la unidad total lo que implica que cada alumno deberá estudiar lo que cada compañero ha explicado. Para Aronson y Patnoe (2011) “although the reward structure is individual rather than group, positive interdependence is built in through task specialization because students need each other to “get the whole picture” (pp. 22–23).

### **1.7.3. *Jigsaw* y sus variantes**

En el tiempo se han diseñado variantes del método *Jigsaw I*. Prueba de ello es la propuesta del *Jigsaw II* por Slavin en 1986, *Jigsaw III* diseñado por Stahl en 1994; *Jigsaw IV* por Holliday en 1995; *Reverse Jigsaw* por Heeden en el 2003 y, el *Subjects Jigsaw* por Doymuş en años recientes, en el 2007. Resulta interesante observar cómo estas variantes comparten en común la subdivisión del tema pero a la vez plantean diferencias sutiles (ver la tabla 7):

Tabla 7: Comparativa entre los métodos *Jigsaw I, II, III y IV*

Características	Métodos
<ul style="list-style-type: none"> <li>– El cuestionario de expertos es asignado a los expertos de cada grupo</li> <li>– Los grupos de expertos contestan los cuestionarios temáticos antes de volver a su grupo</li> <li>– Los estudiantes vuelven a su grupo principal y comparten la información obtenida con sus colegas expertos</li> </ul>	Jigsaw I,II,III y IV
Se revisa todo el proceso	Jigsaw III
Se designan tareas individuales	Jigsaw II,III y IV
<ul style="list-style-type: none"> <li>– Se revisa todo el proceso pero se usa un cuenco con preguntas para todos los estudiantes</li> <li>– Pruebas cortas de los materiales en los grupos de expertos por aspectos temáticos IV</li> <li>– Pruebas cortas sobre el material compartido IV</li> <li>– Se aplica enseñanza por parte del profesor de los temas que hayan quedado incompletos en caso de ser necesario</li> </ul>	Jigsaw IV

Fuente: adaptación de Holliday (2002, p. 2) y Jansoon, Somsook & Coll (2008, p. 182)

La diferencia entre el *Jigsaw I* y el *Jigsaw II*, es que en el segundo, los subgrupos temáticos son premiados al poder completar con éxito el cuestionario de expertos. En el caso del *Jigsaw III* la recompensa es para aquellos alumnos que sean integradores de otros estudiantes cuya lengua sea extranjera, es decir, busca la integración de grupos bilingües. El *Jigsaw IV* incorpora pruebas en mitad de los procesos de búsqueda de información y premia a los grupos o subgrupos temáticos por los alcances logrados en esas pruebas. En el caso del *Reverse Jigsaw*, en lugar de seguir las pautas del docente, la búsqueda de información pretende desarrollar la comprensión de la gama de interpretaciones de los participantes en un número de tópicos pensados sobre la base de la estructura de participación (Hedeen, 2003 citado por Doymuş, Karacop & Simsek, 2010).

Por último, el *Subjects Jigsaw* incluye tres pasos: el tema se divide para la clase en subtemas, luego cada subtema es asignado a grupos de expertos entre 2 y 6 estudiantes quienes trabajan cooperativamente fuera del aula y finalmente, cada grupo complementa las tareas asignadas y presenta la información aprendida al resto de la clase (Doymuş, Karacop & Simsek, 2010).

#### 1.7.4. Dyadic Methods

Aprender en grupos tiene sus encantos, pero hacerlo en pareja tiene ventajas que son descritas por los teóricos de la enseñanza. Como son dos, que deben adoptar conclusiones sobre el aprendizaje, las relaciones se marcan por diferentes condiciones

de acuerdo a los intereses de los profesores y las herramientas que se asuman en el proceso educativo de esa pareja de aprendizaje.

En ese sentido Slavin (1995) ideó los primeros pasos de esta modalidad de enseñanza y, en su planteamiento legó la estructura de los métodos diádicos como una posibilidad distintiva en la que la mayoría de las estrategias cooperativas involucran a cuatro miembros. Dice este autor que existe una larga tradición de estudios de laboratorio que muestran y describen el aprendizaje por parejas, en donde los estudiantes adoptan alternativamente el rol de profesor o alumno por turnos, o bien extraen información conjunta de un texto, experiencias todas que pueden ayudar a incrementar el aprendizaje de los estudiantes “the peer group as the single most powerful influence in undergraduate education” (Falchikov, 2001, p. 1).

Los métodos en parejas se enmarcan en un gran paraguas bajo el nombre de *Peer Tutoring* que cobija diversidad de variantes como *Peer Collaboration*, *Reciprocal Peer Tutoring*, *Scripted Cooperation*, *Peer Teaching*, *Think–Pair–Share*, *Peer Assessment*, *Peer Learning*, *Class Peer Tutoring*, *Guided Peer Questioning*, *Peer Assisted Learning Strategies*, entre otros, siendo los tres primeros los que se han trabajado con mayor frecuencia con estudiantes universitarios y a la vez, los que reportan el mayor número de investigaciones (Johnson & Johnson, 2002a; Johnson, Johnson & Smith, 2007; O'Donnell, 1999).

Entonces ¿en qué se diferencian los términos *Peer*, *Peer Group* y *Peer Tutoring*? Si “a peer is someone of the same social standing, while a peer group consists of those of the same status with whom one interacts (Falchikov, 2001, p. 1) *Peer Tutoring*, es el sistema en el que un alumno enseña a otro o le orienta a cómo solucionar un problema (Dioso–Henson, 2012).

Se trata de que la pareja esté conformada por dos miembros heterogéneos, ya que uno de los miembros adopta la figura de profesor, lo que hace que la interacción sea asimétrica, al ser uno de los componentes el director de la actividad, quien conoce el rol por tener mayor conocimiento de los objetivos a alcanzar y porque además comprende el método y sabe cómo desempeñar su papel. Implica un diseño detallado de las sesiones por parte del profesor (incluyendo los modos de evaluación), el conocimiento necesario de los miembros de la pareja para adoptar los roles, el plan de las sesiones de trabajo en las que el profesor adopta entonces el rol de observador y, por supuesto, la

revisión de lo aprendido por parte de la pareja y la evaluación de lo alcanzado de forma conjunta (Gavilán y Alario, 2010; León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

La pregunta que surge en este método, tratándose de aprendizaje cooperativo en el que se supone todos los miembros ganan ¿En qué se beneficia el profesor? Se beneficia en que al adoptar el rol de tutor, el alumno busca esforzarse para aprender los contenidos al ser responsable de su propia formación y la de su compañero (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005)

Peer tutoring (PT) is characterized by specific role-taking as tutor or tutee, with high focus on curriculum content and usually also on clear procedures for interaction, in which participants receive generic and/or specific training. Some peer tutoring methods scaffold the interaction with structured materials, while others prescribe structured interactive behaviours that can be effectively applied to any materials of interest. Confusion between “tutoring” and “mentoring” is evident in the literature. Mentoring can be defined as an encouraging and supportive one-to-one relationship with a more experienced worker (who is not a line manager) in a joint area of interest. It is characterized by positive role modelling, promotion of raised aspirations, positive reinforcement, open-ended counselling, and joint problem-solving. It is often cross-age, always fixed-role, quite often cross-institution, and often targeted to disadvantaged groups (Topping, 2005, p. 632).

#### 1.7.4.1. *Peer Collaboration*

En el método *Peer Collaboration* –en algunos casos asociados con la tecnología– existe un alto grado de colaboración en cada una de las fases del proceso de aprendizaje, una igualdad en materia de autoridad y propiedad sobre el texto de aprendizaje ya que ningún estudiante se percibe a sí mismo como detentor del conocimiento. La idea es que la pareja trabaje de cerca, de manera conjunta en todas las etapas del proceso con un alto grado de interacción, de igualitarismo y mutualidad (Damon & Phelps, 1989) a través de la comunicación sincrónica y asincrónica.

Si la colaboración “is usually seen as involving comparable effort from all participants towards an outcome from which all benefit equally, but this is not characteristic of all forms of peer tutoring” (Falchikov, 2001, p. 229) ¿cómo comprender el *Peer Collaboration* desde el *Peer Tutoring*? La solución a esta disyuntiva se resume en la siguiente frase “rather than viewing all peer tutoring as a collaborative venture *per se*, it may be more fruitful to consider collaborative technology as a tool for facilitating a variety of peer-tutoring techniques” (Falchikov, 2001, p. 229).

En consecuencia, existen importantes distinciones entre los distintos métodos cooperativos y el *Peer Tutoring* y el *Peer Collaboration*, a pesar de que estas dos técnicas de aprendizaje estén imbuidas bajo el paraguas de la cooperación (Damon y Phelps, 1989) ¿Por qué? porque variará el tamaño de los grupos en función de la técnica aplicada, la contabilidad individual de cada uno de los miembros del grupo en materia de calificaciones, el uso de estructuras externas al propio grupo las tareas alcanzadas y las competencias que se establezcan dentro de los grupos (Slavin, 1983) (ver tabla 8):

Tabla 8: Diferencias entre *Peer Tutoring* y *Peer Collaboration*

		Peer Tutoring	Peer Collaboration
Dimensiones de las parejas	Rol	Bajo en igualdad	Alto en igualdad
	Estructura interactiva	Bajo en mutualidad	Alto en mutualidad
	Tarea	Independiente	Tarea en conjunto
Características sociales de la teoría constructivista	Concepciones de conocimiento	Transmisión universal	Transformación bidireccional
	Zona de desarrollo proximal	El tutor ayuda al tutelado en la transferencia de control	No se enfoca sobre controlar la transferencia
	Discurso	Unidireccional	Bidireccional

Fuente: adaptación de McCarthy & McMahon (1995, pp. 21 y 29).

#### 1.7.4.2. *Reciprocal Peer Tutoring (RPT)*

El método *Reciprocal Peer Tutoring* desarrollado por Fantuzzo y colaboradores se creó inicialmente para ser trabajado en Educación Primaria (Fantuzzo, King & Heller, 1992). Los resultados positivos en el rendimiento académico y otras variables sociales animaron a sus creadores a adaptarlo para ser aplicado con estudiantes universitarios siendo hoy en día el método diádico preferido (Falchikov, 2001) en donde las parejas son similares en cuanto al nivel de rendimiento. Cada miembro alterna el rol de tutor y tutelado en donde el tutor ofrece indicaciones y problemas alternativos para usar frente a su tutelado en caso de que este cometa errores (Slavin, 1995). El proceso favorece la instrucción, evaluación y el refuerzo, lo que conduce a un proceso de asistencia mutua y apoyo social entre los estudiantes.

En la universidad, los estudiantes son asignados a díadas que trabajan durante todo el semestre. Esta dinámica no solo se aplica en el aula de clase sino que se espera que los alumnos continúen trabajando fuera de ella. Finalmente, el día del examen, los alumnos se presentan con la prueba que ha sido trabajada de manera cooperativa dentro y fuera del aula de clase como evaluación final (Fantuzzo, Dimeff & Fox, 1989; Fantuzzo, Riggio, Connelly & Dimeff, 1989; Riggio, Fantuzzo, Connelly & Dimeff, 1991).



Entre otras ventajas se señalan cómo este método ayuda a mejorar los resultados en el ámbito social de los estudiantes (Dufrene, Noell, Gilbertson & Duhon, 2005) y, en especial de los alumnos universitarios (Fantuzzo, Dimeff & Fox, 1989; Fantuzzo, Riggio, Connelly & Dimeff, 1989) porque: (a) facilita la construcción conjunta de conocimientos, (b) se rige por el principio de dar y recibir, (c) la tutoría se convierte en colaboración porque se promueve la multidireccionalidad y (d) reduce el autoritarismo y la dependencia entre tutor–tutelado porque ambos miembros de la pareja se benefician de una relación simétrica (Duran y Vidal, 2004). Reconoce que la metacognición es un mediador importante para el éxito de la enseñanza en Educación Superior, ya que la mayoría de los estudiantes universitarios no cuenta con herramientas para regular adecuadamente su autoaprendizaje (De Backer, Van Keer & Valcke, 2012).

#### *1.7.4.3. Scripted Cooperation (SC)*

Los problemas inherentes a intentar analizar los mecanismos subyacentes en un contexto fluido de instrucción regular poseen serios problemas de tratamiento para efectos de validación de la investigación en aprendizaje cooperativo. En consecuencia, los estudios de Dansereau, O'Donnell y Lambiotte (1988 citado por O'Donnell & Dansereau, 1995) se enfocaron en desarrollar un contexto controlado de experimentación: una estrategia cooperativa pensada en la escritura entre pares. Así surge el *Scripted Cooperation* dentro de un laboratorio que envuelve a los estudiantes en un trabajo en conjunto estructurado en díadas para dominar una tarea académica o técnica siendo este método usado como que si de una representación teatral se tratara.

Incluye cinco pasos que facilitan el aprendizaje: (1) dividir el texto en secciones discretas y significativas; (2) que los dos miembros de la díada lean el texto a la vez; (3) requiere que uno de los miembros recuerde los aspectos y detalles de la información pertinentes (esta función debe ser alternativa entre los dos miembros de la pareja de estudio); (4) requiere que el otro, el que escucha a quien relata los temas principales, detecte los errores y omisiones de su compañero; y, (5) exige que los dos miembros de la díada elaboren su información con métodos que pueden incluir el desarrollo de analogías y la creación de imágenes (Anderson, 1985 citado por Hertz–Lazarowitz, Benveniste & Miller, 1995). Este método también fomenta la metacognición y la transferencia de contenidos desde una forma de posesión o dominio de aprendizaje hacia otra en distintas asignaturas.

O'Donnell y Dansereau (1995) consideran que los contenidos independientemente de los resultados afectivos, simultáneamente desvelan un resultado del aprendizaje por observación durante la interacción con guión: cada miembro de la díada debe brindar a la interacción colaborativa su repertorio de conocimientos y herramientas efectivas que podría haber adquirido desde el modelaje social y las fortalezas ensayadas.

#### *1.7.4.4. Peer Teaching*

A diferencia de los métodos *Scripted Cooperation* y *Reciprocal Peer Tutoring* que han sido categorizados como técnicas del *Peer Tutoring* mas complejas, en el *Peer Teaching*, como su nombre lo indica, los estudiantes asumen el rol de profesor. Así lo concibieron Goldschmid y Goldschmid en 1976, en donde los estudiantes de similares edades o nivel educativo enseñan a otros. Es precisamente en la acción en donde se asume el rol del profesor “He is forced to transform knowledge (facts, concepts) in such a way that it may be assimilated by his partner (...) the most profitable peer teaching situation would be one where roles of “teacher” and “student” are alternated” (p. 13).

En la universidad, este método refuerza la autonomía y recuerda la responsabilidad en el propio aprendizaje, incrementa la autoestima y fomenta una actitud positiva hacia el aprendizaje. La técnica requiere que el estudiante que asuma el rol del profesor desarrolle competencias de búsqueda de información, elabore preguntas apropiadas, ofrezca una retroalimentación pertinente y establezca una relación que conduzca al aprendizaje. Como cualquier otro método, los creadores advierten que no puede reemplazar otras formas de instrucción, sino que debería trabajarse en conjunto con otras técnicas de aprendizaje, enriqueciendo así la experiencia.

De esta forma, homogeneidad o heterogeneidad serán la marca distintiva que tengan los métodos en parejas. La relación con el profesor, el estilo de tutorías, las posibilidades de intercambio de conocimiento, son y serán los signos de un proceso que, en díada, generará conocimientos y modos de enfrentar los problemas y tareas de aprendizaje propios del ámbito universitario.

#### *1.7.4.5. Think-Pair-Share*

*Think-Pair-Share* es una técnica clasificada como sencilla en la dinámica de las parejas cooperativas. Sherman (1991) atribuye el descubrimiento de este método a Lyman (1981) en su capítulo *The responsive classroom discussion: The inclusion of all students*

en el libro compilado de Anderson titulado *Mainstreaming Digest* aunque resulta ser una técnica muy común en la docencia.

Este método se produce cuando un docente presenta el tema a la clase y los estudiantes se sientan en parejas que pertenecen a su grupo; entonces, el docente plantea preguntas a la clase y los alumnos deben pensar las respuestas por ellos mismos, en pares, adoptando una respuesta consensuada. Al final, el profesor hará preguntas para llegar a una respuesta única (Kagan, 1994).

Es útil para desarrollar la confianza de los estudiantes al permitir que las parejas “ensayen” una respuesta que ha sido discutida y acordada para luego compartirla con toda la clase. También se aplica para la reflexión de una lectura de un texto siendo una oportunidad que los alumnos compartan sus apuntes e ideas (Falchikov, 2001).

#### **1.7.5. *Learning Together* (LT)**

Johnson y Johnson idearon, dentro del espectro de posibilidades del aprendizaje cooperativo, un método que requiere de una formación y preparación sólida de los profesores para su implementación en el aula es el *Learning Together*, pero autores como Santos Rego (1990) considera que se no se trata de un método cooperativo sino de una filosofía de la cooperación.

Esta estrategia incluye los siguientes elementos: (a) tres tipos de procedimientos de aprendizaje cooperativo usados de una manera integrada como lo son el aprendizaje cooperativo formal, aprendizaje cooperativo informal y grupos cooperativos de base, (b) cada lección o actividad está conformada por los principios esenciales que guían el trabajo cooperativo, (c) una rutina cooperativa consolidada y (d) una estructura organizacional del centro educativo que apoya el trabajo en grupo (Johnson & Johnson, 1989; Johnson, Johnson & Holubec, 1992, 1999; Johnson, Johnson & Smith, 1991 citados por Johnson & Johnson, 1999a).

Los profesores usan el aprendizaje cooperativo de modo suficiente para alcanzar un uso rutinario de competencia y una participación regular entre los colegas que sustentan a los grupos de participación dentro de la toma de decisiones de la institución educativa (Johnson & Johnson, 1999a). Los pasos necesarios para una adecuada implementación son similares al STAD aunque también destaca la construcción y la autoevaluación grupal: (a) seleccionar el tema, (b) determinar el tamaño del grupo, (c) asignar a los estudiantes a grupos heterogéneos de trabajo, (d) adecuar la clase para promover

interacciones cooperativas entre estudiantes, (e) seleccionar materiales y recursos que promuevan la interdependencia positiva, (f) orientación y seguimiento por parte del profesor a cada uno de los grupos y (g) una evaluación individual y en donde la media del equipo será la calificación final del grupo (Johnson & Johnson, 2014a).

En ese sentido, para hacer efectivo al aprendizaje cooperativo, los profesores comprenden que la naturaleza de la cooperación y la esencia de los componentes son indispensables en una bien estructurada lección cooperativa. Entender cuán positiva es la interdependencia, la promoción de interacción, la responsabilidad individual, las herramientas sociales, el procesamiento de grupo y las herramientas para el desarrollo en estructuras, hace que los profesores deban adaptar el aprendizaje cooperativo a estas circunstancias únicas con el fin de afinar el uso del aprendizaje cooperativo, y así los estudiantes trabajen juntos en la solución de problemas que se les plantean en el aula (Johnson & Johnson, 1999a).

El reto del *Learning Together* es la complejidad que comporta su instauración porque requiere de “una cultura y política que arrope a la institución educativa”, que sea transversal a todo el proceso de formación y que sea compartida y aplicada por todos: profesores, directivos y, por supuesto, estudiantes.

#### **1.7.6. Informal Methods**

Existen algunos métodos que bajo el título de “informales” agrupan fórmulas muy particulares de aprendizaje cooperativo. Estos métodos diseñados para ser trabajados desde la etapa de primaria hasta la universidad, se caracterizan por ser breves y pueden utilizarse bien para iniciar o cerrar una actividad, para presentar un tema en una clase tradicional o para combinarse con métodos cooperativos como el STAD, TGT u otros. Es Slavin (1995) quien escribe una nota al margen sugiriendo que su clasificación es una adaptación de la propuesta de Kagan (1992):

*Spontaneous Group Discussion*: ocurre cuando los estudiantes están sentados en grupos y en el momento de la lectura de la clase se produce la necesidad de que resuelvan alguna inquietud propia de la asignatura que se está impartiendo.

*Numbered Heads Together*: el método surgió cuando Russ Frank, un profesor de educación media dijo a sus estudiantes que pusieran sus cabezas juntas para resolver la ubicación de una coma en un texto. Esa animada discusión que se produce en el aula es

la que dio pie a este método. Es básicamente una variante del grupo de discusión que tiene como factor de interés el respaldar la responsabilidad individual.

*Team Product*: los estudiantes hacen un centro de estudio, es decir, escriben un ensayo, trabajan en una presentación, etcétera. Es importante en este método que las tareas estén muy bien delimitadas para que la responsabilidad individual pueda tener claros sus objetivos.

*Cooperative Review*: los alumnos se preparan en conjunto para un examen. Puede ser que el profesor haya dado las preguntas del examen y eso permita que el grupo tenga las respuestas correctas y, bajo los criterios de recompensa algún otro grupo disponga de elementos que agregar a esa respuesta. Puede ser que se distribuyan en el aula más de dos grupos y que sobre la dinámica de respuestas y ampliaciones de respuestas se desarrolle la clase de revisión cooperativa.

El procedimiento a seguir en la mayoría de estos métodos informales es el siguiente: (a) el profesor elabora una pregunta antes de la actividad planificada; (b) el grupo se reúne a discutir; (c) se lleva a cabo la actividad planificada que puede ser un video, una lectura o demostración; (d) se detiene tantas veces la actividad como sea necesario para que el grupo se reúna a discutir; (e) el profesor elabora otra pregunta que ayuda a los estudiantes a sintetizar y resumir el material que ha sido presentado y, por último; (f) el grupo se reúne nuevamente para discutir lo que han aprendido (Ornstein & Scarpaci, 2012).

Para Kagan (1990) “structures differ in their usefulness in the academic, cognitive, and social domains, as well as in their usefulness in different steps of a lesson plan” (p. 12). Lo más importante es que el profesor se pregunte qué competencias cognitivas, académicas y sociales pretende desarrollar para así encontrar aquella estructura cooperativa que favorezca el aprendizaje y en consecuencia, incrementar el rendimiento académico.

### **1.7.7. Cooperative Learning and Computers**

A partir de la irrupción de las Tecnologías de Información y Comunicación las posibilidades de interacción de los grupos tienen nuevos escenarios de colaboración e interacción. La diversidad de métodos colaborativos es enorme “Collaborative learning research suggests new models of schooling in the 21st century by new technologies” (Zhang, 2013, p. 502). Lo interesante, es que asumen las premisas y principios de la

colaboración cara–cara y las adaptan a la tecnología ofreciendo también resultados en el rendimiento académico y en las habilidades sociales (Kulik & Kulik, 1986; Kulik & Kulik, 1991; Kulik, Kulik & Cohen, 1980; Male, 1994).

Del método *Computer Supported Collaborative Learning* (CSCL) se han desarrollado otros como *Scripts in Computer–Supported Collaborative Learning* o *Mobile Computer Supported Collaborative Learning* (MCSCCL o mCSCL) al que se suman métodos que incorporan el aprendizaje colaborativo en redes como *Learning through Collaborative Designing* (LCD) o aquellos que enfatizan el uso de la tecnología como apoyo en el trabajo *Computer Supported Collaborative Work* (CSCW), en comunidades de aprendizaje como *Fostering a Communities of Learners Model* (FCL) o *Knowledge Building Community Model* (KBC) y otros mas específicos como *Computer–Assisted Instruction* (CAI) y *Collaborative Knowledge Construction* (CKC), entre otros.

Todos estos modelos, sin duda se apoyan en la interacción hombre–máquina y la disparidad resalta en los pequeños matices que las propias herramientas tecnológicas proponen o prometen. Los próximos métodos colaborativos son un pequeño ejemplo de los más utilizados en la universidad.

El método *Computer–Assisted Instruction* (CAI) creado en la década de los 60 del siglo pasado pretendía ser un método de enseñanza independiente. Su valor era el estudiar al propio ritmo del estudiante en programas como matemáticas o gramática donde las respuestas eran ciertas o falsas. De esa manera, la fortaleza del CAI está en que los resultados de aprendizaje rondan el 20% o más pero, como contrapartida tiene dos elementos poco favorables: la obsolescencia del software, y que se quedó como una promesa más que revolucionaría el aprendizaje humano (Ehrmann, 1995; Kulik, Chen–Lin & Kulik, 1991 y citados por Millis & Cottell, 1998).

Esta fallida oferta encaminó a los educadores a buscar maneras de reducir la soledad de los educandos y los costes de formación a través de los ordenadores. Surgió entonces el segundo método, el *Computer–Assisted Learning* (CAL). Así, el foco de la atención se centró en dos factores: lo informático y lo cooperativo (Light & Mevarech, 1992 citados por Millis & Cottell, 1998) y la investigación se ubicó en conocer las múltiples variables influyentes en el proceso de enseñanza–aprendizaje: diseño de tareas, tamaños y dinámicas de los grupos de aprendizaje, edades, sexo y tiempo dedicado a la tarea, entre otros.

Resulta interesante comprender cómo motivos económicos y académicos condujeron a la creación de métodos colaborativos. Afirman Rysavy y Sales (1991) que los altos costos en materiales y equipos tecnológicos, el acceso limitado a los ordenadores para la enseñanza en muchos centros educativos, y el éxito de las estrategias de aprendizaje cooperativo en comparación con la enseñanza tradicional en el aula, fueron algunas razones que llevaron a los investigadores a examinar el potencial del método *Cooperative Based Instruction* (CBI), el cual estaba diseñado para trabajar de manera individual y no para el aprendizaje colaborativo.

Hooper (1992) quien compartió el argumento de sus colegas hizo en su momento un llamado a cambiar la manera de concebir los softwares “the nature of instructional strategies for individualized CBI may be inappropriate for group learning” (p. 22). Los resultados de este método colaborativo fueron controversiales porque algunas investigaciones obtuvieron resultados positivos en el rendimiento cuando los estudiantes trabajaban colaborativamente en comparación con el trabajo individual y en otros estudios, no se obtuvieron diferencias significativas (Carrier & Sales, 1987).

Por su parte, el método *Computer-Assisted Language Learning* (CALL) surgió en la década de los años 90 con la intención de favorecer el aprendizaje de una segunda lengua (Warschauer, 1996; Warschauer & Healey, 1998). Este método se aplicó para el desarrollo de habilidades cooperativas comunicativas entre pares en situaciones reales. Numerosos estudios siguen confirmando el uso de este método como una herramienta que favorece un aprendizaje cooperativo entre nativos y aprendices de una lengua extranjera al incrementar las habilidades de leer, escuchar, hablar y escribir (Mutlu & Eroztuga, 2013).

Koschmann, es quien en el año 1996, emplea por primera vez el nombre de “aprendizaje colaborativo mediado por ordenadores” en inglés *Computer Supported Collaborative Learning* (CSCL) que es uno de los métodos mas utilizado y del que se ha estudiado con mayor rigurosidad (Stahl, Koschmann & Suthers, 2006). Se caracteriza por la unión entre teoría e investigación en aulas de clase con aprendizaje cooperativo y la mediación de comunicación a través de ordenadores (CMC) con el propósito de comprender cómo los grupos basados en CMC pueden colaborar para aprender. Los beneficios del uso del CSCL están en el incremento de la responsabilidad estudiantil, iniciativa, participación y aprendizaje (Dennen & Hoadley, 2013) en donde

intervienen factores como la tarea, el medio de comunicación, el tutor y el estudiante (Veerman & Veldhuis–Diermanse, 2006).

Se trata de un método destinado a los niveles superiores de enseñanza para incentivar la comunicación y discusión conceptual entre pares (Brandon & Holingshead, 1999 y Feather, 1999 citados por McWhaw, Schnackenberg, Sclater & Abrami, 2003). Una de las características principales del CSCL es el uso de las habilidades sociales siendo necesario que el docente diseñe actividades para involucrar a los participantes (Abrami et. al, 1995; Towns, 1998 citados por McWhaw, Schnackenberg, Sclater & Abrami, 2003). Las prácticas de los docentes requieren fundamentalmente diferentes fuentes, comunicación con los colegas de profesión, así como otras maneras de comunicación on–line. De esa manera lo importante es tratar de que el profesor pueda adecuar su labor a las nuevas tecnologías para poder brindar a sus estudiantes el apoyo necesario en aprendizaje cooperativo a través de ordenadores (Williams & Kelly, 2006).

Por último, otro de los modelos que se combina frecuentemente con métodos colaborativos que usan a la vez la tecnología, es el *Problem–Based Learning* (PBL), el cual pretende facilitar el proceso de investigación en grupos cooperativos, ya que los estudiantes tienen que resolver auténticos problemas con el fin de promover herramientas del aprendizaje autorregulado, razonamiento y comunicación (O'Donnell & Hmelo–Silver, 2013; Savin–Baden, 2000).

### **1.7.8. Los métodos cooperativos y sus diferencias**

Entre los métodos de aprendizaje cooperativo existen diferencias relativas en el sistema de recompensas, la especialización de las tareas, las competencias entre los grupos, la igualdad de oportunidades y las tareas y materias académicas.

La labor que se pretende en este epígrafe es precisamente destacar las similitudes y disparidades entre los métodos STAD, TGT, TAI, LT, CIRC, GI, *Jigsaw I*, *Jigsaw II* según las características mencionadas.

Como se puede apreciar en la tabla 9, en el ámbito del *sistema de recompensas* el aprendizaje cooperativo pretende que los alumnos trabajen en grupo, en conjunto, para alcanzar objetivos de aprendizaje, pero el enfoque de la evaluación variará de acuerdo a los métodos e incluso, a los autores que lo traten. Las recompensas de grupo se pueden



sustentar sobre la base del (a) aprendizaje individual, (b) el producto grupal o (c) simplemente sin ninguna gratificación.

Tabla 9: Clasificación métodos cooperativos según estructura de recompensa

Recompensas de Grupo, basadas en los aprendizajes individuales	Recompensas de Grupo, basadas en el producto grupal	Sin recompensas de Grupo
STAD, TGT, TAI, CIRC, JIGSAW II	LT	GI y JIGSAW I

Fuente: León del Barco, Gozalo, Felipe, Gómez y Latas (2005, p. 79)

Así, en los métodos de Slavin como STAD, TGT, TAI, CIRC y *Jigsaw II* (c.p. Santos Rego, 1990 en León del Barco, Gozalo, Felipe, Gómez y Latas, 2005) es importante el control y la evaluación individual, mientras que para los métodos *Jigsaw I*, GI y LT el valor fundamental está en la capacidad de cooperación entre compañeros de grupo para alcanzar los resultados del aprendizaje.

En consecuencia “... las recompensas de grupo deben basarse en el aprendizaje de todos los miembros para que den como resultado un aumento en el aprendizaje de cada uno de ellos” (Gavilán y Alario, 2010, p. 203), lo que supera con creces al propio contenido. La experiencia está entonces en poder demostrar de manera individual lo que se ha adquirido dentro del grupo de aprendizaje.

Por otro lado, dentro del mismo sistema de recompensas, los métodos STAD, TGT, TAI, CIRC y *Jigsaw II* están estructurados sobre las tareas de grupo buscando destacar que las recompensas grupales sólo se obtienen cuando cada uno de los miembros es capaz de aportar al grupo porque “...each team member receives an individual score that contributes to the team score, so the reward structure is probably best describes as individualistic and cooperative” (Kagan, 1985, p. 85). Y, LT considera que las recompensas grupales sólo se obtienen en función del producto que se logre en conjunto y no por los aprendizajes individuales (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

Finalmente *Jigsaw I* y GI no ofrecen ninguna recompensa al grupo una vez finalizada la tarea aunque ambos métodos enfatizan los procesos grupales, por consiguiente, la eficacia en el rendimiento de los estudiantes disminuye al compararse con otros métodos tradicionales.

Ahora bien, cuando se compara este sistema de recompensas con otros factores los métodos se diferencian aun mas. Es León del Barco y sus colegas (2005) quienes

elaboran una propuesta interesante a través de un cuadro comparativo que permite visualizar fácilmente estas disimilitudes (ver tabla 10)

Tabla 10: Clasificación métodos cooperativos según estructura de recompensas, tareas e igualdad de oportunidades para la puntuación

Factores		Recompensas de Grupo, basadas en los aprendizajes individuales	Recompensas de Grupo, basadas en el producto grupal	Sin recompensas de Grupo
Especialización	Sin especialización de tareas	STAD, TGT, TAI, CIRC	LT	
	Con especialización de tareas	JIGSAW II		GI y JIGSAW I
Competición	Sin competición entre grupos	TAI, CIRC	LT	GI y JIGSAW I
	Compitiendo entre grupos	STAD, TGT, JIGSAW II		
Igualdad oportunidades	Sin igualdad de oportunidades		LT	GI y JIGSAW I
	Con igualdad de oportunidades	STAD, TGT, JIGSAW II, TAI, CIRC		
Materias	Para todas las tareas y materias	STAD, TGT	LT	
	Tareas y materias específicas	JIGSAW II (Ciencias Sociales) TAI, (Matemáticas) CIRC, (Lectura y escritura)		GI (Realizar proyectos) JIGSAW I (Ciencias Sociales)

Fuente: adaptación de León del Barco, Gozalo, Felipe, Gómez y Latas (2005).

Según estos autores, si se ubica la mirada sobre la *especialización de la tarea*, hay coincidencia en que el objetivo es alcanzar –en los métodos *Jigsaw I*, *Jigsaw II* y GI– la interdependencia entre los miembros del grupo, que se logre la tarea gracias al esfuerzo de todos y cada uno de los miembros, debido a que se ha dado a cada uno de ellos un ámbito de responsabilidad particular, lo que evita que se difuminen los roles (Gavilán y Alario, 2010).

A pesar de que los demás métodos –STAD, TGT, TAI y CIRC– no consideran importante la especialización de la tarea, la asignación de roles interconectados es importante para los métodos TAI, CIRC y LT.

El propio aprendizaje cooperativo se debate en sus métodos entre poner a competir a los grupos entre sí, o no hacerlo. En este sentido, están a favor de esta *competencia entre grupos* –con miras a motivar la participación y la cooperación– los métodos STAD,

TGT y *Jigsaw II*. Todo lo contrario proponen los métodos *Jigsaw I*, LT, GI, TAI y CIRC (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

En el mismo sentido, Slavin (1985) citado por León del Barco, Gozalo, Felipe, Gómez y Latas (2005) considera que la “competencia es motivadora para los alumnos y aceptable también para una amplia gama de maestros” (p. 80), aumentando el rendimiento en situaciones de aprendizaje cooperativo. Por su parte, Johnson, Johnson, Holubec y Roy (1984) –citados también por estos mismos autores– se transan por la cooperación intergrupal en lugar de la competencia intergrupal.

Los métodos de aprendizaje cooperativo además plantean diferencias en cuanto a la *igualdad de oportunidades para la puntuación*. Así, en los métodos STAD, TGT, TAI, CIRC y *Jigsaw II*, si los miembros del grupo hacen su máximo esfuerzo, estos pueden contribuir al puntaje del equipo, aumentando la motivación hacia el éxito de todos los integrantes del grupo, y no solo para aquellos que sean más aventajados para los procesos de aprendizaje. Asimismo, estas posibilidades permiten compartir la responsabilidad. El resto de los métodos de aprendizaje cooperativo –GI y *Jigsaw I*– no consideran como propio la igualdad de oportunidades, lo que se traduce en diferentes calificaciones para cada uno de los miembros del grupo (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005).

La gran idea que subyace en cuanto a igualdad de oportunidades es que todos y cada uno de los miembros de un equipo sienta que es útil su contribución al aprendizaje cooperativo, que por el hecho de participar y oponer su aprendizaje frente a los otros, tiene las mismas oportunidades de ser valorado por su progreso, por su esfuerzo individual, lo que hace que los menos aventajados se sientan partícipes al tener que superarse y poder aportar al grupo sus puntos (Gavilán y Alario, 2010).

Por otra parte, los diferentes métodos de aprendizaje cooperativo son útiles en todas las *tareas o disciplinas académicas* pero sin olvidar su especificidad. En ese sentido, si se pretende enseñar Ciencias Sociales, *Jigsaw I* y *Jigsaw II* son métodos ideales; TAI para las Matemáticas; CIRC para el afianzamiento de la Lectura y Escritura en los primeros niveles de formación; STAD, TGT y LT son mucho más dúctiles en todas las disciplinas. Y, si de niveles de formación se habla, TAI y CIRC son muy usados en primaria, STAD y TGT en primaria y secundaria, LT y GI son más adecuados en

secundaria (León del Barco, Gozalo, Felipe, Gómez y Latas, 2005) aunque muchos se han llevado a cabo también en bachillerato y en la universidad.

Como muy bien dicen Johnson, Johnson y Stanne (2000) “Knowing that cooperative learning can significantly increase student achievement (compared with competitive and individualistic learning) when properly implemented does not mean, however, that all operationalizations of cooperative learning will be effective or that all operationalizations will be equally effective” (s.p.). Lo importante es que los profesores adquieran competencias para el diseño de métodos cooperativos e incorporen al currículo metodologías de enseñanza que incrementen el rendimiento académico de los alumnos.

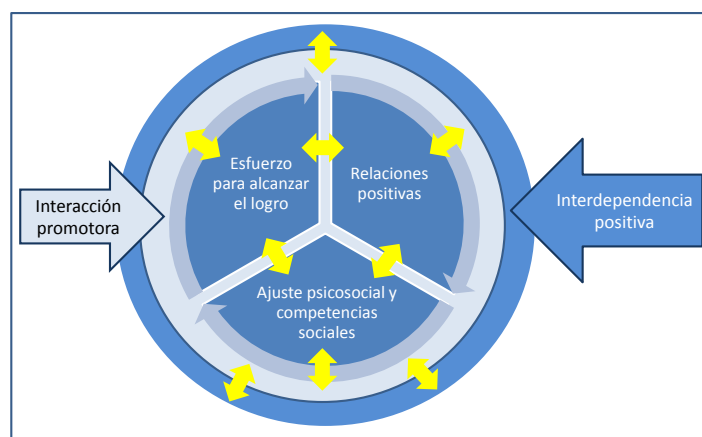


## Capítulo II. APRENDIZAJE COOPERATIVO: EVIDENCIAS EMPÍRICAS

### 2.1. Resultados del aprendizaje cooperativo

La eficacia del aprendizaje cooperativo se ha medido en tres resultados amplios e interrelacionados entre sí: (a) el esfuerzo ejercido para alcanzar el logro, (b) relaciones positivas y (c) el ajuste psicosocial y las competencias sociales. Estos resultados son consecuencia de la relación recíproca entre la interdependencia positiva y la promoción de la interacción entre participantes (Johnson & Johnson, 1989; Johnson & Johnson, 1999b; Johnson, Johnson & Smith, 1991) (ver ilustración 2).

Ilustración 2: Resultados del aprendizaje cooperativo y patrones de interacción



Fuente: Johnson, Johnson & Smith (1991, p. 29)

En la dimensión *relaciones positivas* las investigaciones se han orientado a estudiar (a) el apoyo social, (b) el abandono escolar, (c) las relaciones de la facultad con los estudiantes, (d) las relaciones interpersonales, (e) las relaciones entre pares, (f) la multiculturalidad en el aula y (g) la integración social, tiempo, conducta y asistencia a clase (Slavin, 1995).

A este inventario, se suman otras variables que han sido medidas bajo la categoría *ajuste psicológico y las competencias sociales* como son (a) la autoestima, (b) las habilidades sociales; (c) el altruismo, las actitudes constructivas y el punto de vista del otro y (d) el locus de control interno (Slavin, 1995).

El otro resultado que también ha sido centro de interés en la comunidad científica con una larga tradición en Psicología Social y Psicología Educativa, es el rendimiento y la productividad de los grupos cooperativos. Es una de las variables que conforma la dimensión *esfuerzos para alcanzar el logro*, en conjunto con (a) la calidad de las estrategias que se utilizan, (b) las ganancias que ocurren en el proceso cuando nuevas ideas, soluciones o esfuerzos son generados o creados y (c) la transferencia del aprendizaje de conocimientos, estrategias o habilidades (Johnson & Johnson, 1989).

Es el mismo Slavin (1980) quien afirma que medir el rendimiento es siempre una tarea difícil en lo sustantivo –definición conceptual– y en lo metodológico –definición operacional–

Many studies used curriculum-specific tests to measure achievement. This is legitimate only when the content and rate of progress in experimental and control groups is held strictly constant; otherwise, students could be learning things other than what they were tested on to different degrees in the different treatments. Other studies used only standardized tests, which have the problem of being so comprehensive that they are insensitive to change due to treatments. Achievement tests can measure facts and low-level skills, or they may measure high-level understanding of material. Only the Sharan studies make a specific breakdown of treatment effects by the concept level of the test, but lack of concrete understanding of what constitutes "achievement" may explain contradictory outcomes in different studies (p. 325).

Por consiguiente, el extenso volumen acumulado de investigaciones en aprendizaje cooperativo invita a conocer una breve panorámica no exhaustiva en el contexto internacional y nacional con la finalidad de ofrecer una perspectiva general actualizada porque “Cooperative learning, especially, in the last decade, has attracted interest in teaching and research at college level education” (Brewer & Klein, 2006, p. 24).

### **2.1.1. Perspectiva internacional de investigaciones en la universidad**

Resultados y valoraciones positivas sobre el rendimiento en equipo y el aprendizaje colaborativo (Kapp, 2009; Massengill, 2011), la actitud estudiantil hacia el aprendizaje cooperativo (Akhtar, Perveen, Kiran, Rashid & Khatoon, 2012; Campisi & Finn, 2011), la autoevaluación del trabajo en pequeños grupos (Aydin, 2013), los beneficios para el aprendizaje y la retención de lo aprendido en la rendición de exámenes cooperativos (Bloom, 2009) o las comparaciones entre el aprendizaje cooperativo y los métodos tradicionales de enseñanza (Baghcheghi, Reza & Rezaei, 2011; Beck & Chizhik, 2013), son algunos ejemplos de investigaciones recientes en el ámbito universitario.

También destacan investigaciones relativas a los juegos de rol que se aplican en el aprendizaje colaborativo (Poling & Hupp, 2009), cómo las dificultades para ponerse en el lugar del otro dentro de un grupo generan perspectivas diferentes con respecto a los contenidos de aprendizaje (Gillespie & Richardson, 2011) o cómo la puntuación justa de acuerdo a la contribución tanto de carácter individual como en parejas permiten comprender la dinamización de los usos del aprendizaje cooperativo (Cvetkovic, 2013).

Otros estudios han evaluado lo transcultural observando que las diferencias culturales no generan disparidad a la hora de aprender con técnicas colaborativas (Bulut, 2009), el trabajo en pequeños grupos como estrategia que favorece las relaciones interraciales (Ouellett & Fraser, 2011), la influencia del aprendizaje cooperativo en los distintos estilos de estudio para promover un pensamiento crítico, reflexivo y multidimensional (Güneysu & Tekmen, 2010) y, la percepción estudiantil sobre las herramientas de comunicación interpersonal y el éxito profesional luego de haber comprendido y usado métodos cooperativos de aprendizaje (Ballantine & McCourt, 2009).

A este inventario se suman investigaciones orientadas a cómo el conflicto competitivo del manejo de información similar entre los pares de un grupo se disipa cuando esa información es diferenciada, lo que enriquece significativamente los aportes al grupo (Buchs, Pulfrey, Gabarrot & Butera, 2010), los efectos del aprendizaje cooperativo en parejas versus el aprendizaje individual en materia de motivación y recompensas (Sears & Pai, 2012) o una investigación similar que evalúa las relaciones entre el aprendizaje colaborativo, la autonomía en el aprendizaje y la falta de cortesía entre estudiantes universitarios, en los que la atención personalizada del profesor promueve cambios en el proceso de interacción positiva hacia el aprendizaje cooperativo (Summers, Bergin & Cole, 2009) y, ¿por qué no? investigaciones centradas en la formación continua del profesorado en donde la reflexión permanente de la práctica educativa se convierte en el saber y en el hacer, en el conocer y en el convivir (Mansour, EL-Deghaidy, Alshamrani & Aldahmash, 2014 en prensa). De esta manera, Crockrell, Caplow & Donaldson (2000) “reviewed several teaching methods and concluded that cooperative learning is the most promising of teaching tools” (p. 24).

#### *Experiencias cooperativas en diferentes asignaturas*

Para la enseñanza de la Estadística, las actividades cooperativas favorecen el rendimiento cuando se compara con el método tradicional de clase (Rabin & Nutter–



Upham, 2010); permiten reflexionar sobre la efectividad de los procedimientos de los contenidos del curso y gestionar mejoras pedagógicas para que las tareas estudiantiles, faciliten la docencia (Robinson & Cooper, 2010) y reduzcan las actitudes negativas hacia las matemáticas al incrementar los niveles de motivación, al ofrecer que los estudiantes se familiaricen con el trabajo con sus pares y lideren nuevos fundamentos para disfrutar las clases (Panitz, 2010). Cuando se aplican en conjunto con la metodología *Problem Based Learning* (PBL) fomentan la enseñanza de la contabilidad (Cottell, 2010) y con métodos como *Think–Pair–Share*, *Jigsaw* y una controversia constructiva favorecen contenidos de introducción en Economía (Maier, McGoldrick & Simkins, 2010). Y en un primer curso de Administración los beneficios se potencian cuando el grupo es homogéneo en cuanto al sexo y las habilidades académicas y sociales (Opdecam, Everaert, Van Keer & Buysschaert, 2014).

La experiencia de la Química no se queda atrás. El método *Jigsaw* y las técnicas de animación se han utilizado para la comprensión de conceptos relativos a la naturaleza de la materia (Karacop & Doymuş, 2013) y para comprobar cómo el rendimiento mejora en la retención y la cobertura del contenido de la asignatura (Shadle, 2010). En otras ramas del saber universitario, en la Ingeniería, esta metodología se ha aplicado a lo largo de 30 años y ha sido continuamente mejorada por profesores universitarios (Smith, Matusovich, Meyers & Mann, 2010).

Para maximizar la enseñanza de la Ciencia, el trabajo en grupo es fundamental en los experimentos científicos (Schrock, 1995). En Biología, el aprendizaje cooperativo es esencial porque promueve la comprensión de los textos (Nelson, 2010) y favorece que los estudiantes sean capaces de escribir correctamente sus informes e investigaciones (Liao & Worth, 2011). En Veterinaria, Schoeman, Van Schoor, Van der Merwe & Meintjes (2009) utilizaron los grupos pequeños para la comprensión de problemas clínicos. En Geología, las oportunidades de trabajo en grupo son enormes porque probablemente sea la única forma de aprender algunos contenidos de esta ciencia (Nuhfer, 2010).

Otras investigaciones han dado un vuelco a las asignaturas “más duras” y han demostrado las bondades en el ámbito de la Literatura porque el éxito de esta metodología radica en que el estudiante aprende bajo la construcción en clase al ser participativo, interactivo con los materiales y motivador de su proceso de aprender (Millis, 2010a). También el aprendizaje cooperativo se ha aplicado en asignaturas

interdisciplinarios como Música y Arte (Richter & Thomas, 2011) y en el desarrollo de habilidades sociales e interpersonales en clases de Psicología de la Educación bajo la tutela del profesor universitario (Cohen, 2010).

#### *Los métodos cooperativos: diversidad*

Investigaciones recientes han comparado la eficacia de los métodos cooperativos en habilidades cognitivas, sociales y efectivas. El STAD está suficientemente probado por su efectividad en materia educativa universitaria cuando se combina con tecnología a través de mensajes de texto –SMS– (Liu et al. 2005), con el método CIRC para el aprendizaje de una segunda lengua (Zarei, 2012) o para el trabajo en equipo en laboratorios de Biología (Maloof & White, 2005) o cuando se compara a estudiantes de Enfermería que trabajan en grupos competitivos y no competitivos (Wang, 2012).

El método GI tampoco se queda a la cola. Varios estudios se han desarrollado en la universidad con este método caracterizado por una mayor flexibilidad en la asignación de roles y tareas. Temas tan variados como la enseñanza de la Comunicación, Álgebra, democracia en el Medio Este, las percepciones en Psicología y la Geografía demuestran su eficacia y su fácil aplicación en el aula de clase (Fay & Hardie, 2003; Geok–Chin, Kim–Eng & Sharan, 2007; Geok–Chin, Sharan & Kim–Eng, 2005; Huhtala & Coughlin, 1991; Parkay, Oaks & Peters, 2000).

Otros hallazgos en el ámbito universitario dan fe de la utilidad del método *Jigsaw* en todas sus modalidades. Por ejemplo, Jansoon, Somsook & Coll (2008) usaron *Jigsaw IV* para el aprendizaje de experiencias en prácticas de Química; Sedat (2010) aplicó el mismo *Jigsaw IV* en la formación inicial de futuros maestros de lengua; Doymuş, Karacop y Simsek (2010) emplearon *Jigsaw* para la comprensión de conceptos electroquímicos y sus colegas Perkins & Saris (2001) también se valieron del *Jigsaw* para la enseñanza de la Estadística y no por menos importantes, Huang, Huang & Hsieh (2008) combinaron este mismo método con tecnología y Lai & Wu (2006) desarrollaron su investigación en un entorno de aprendizaje cooperativo usando el método *Jigsaw* con *handhelds* inteligentes (ordenadores de mano) conectados inalámbricamente entre estudiantes de Enfermería.

Relativos a este último ámbito, los métodos cooperativos que se sustentan en el uso de los ordenadores y las tecnologías de información y comunicación se resumen en estudios interdisciplinarios en donde la transversalidad del conocimiento es una de las

demostraciones más fehacientes del uso de esta metodología en la educación universitaria (Dillenbourg, 2003). Estas investigaciones versan en torno a la enseñanza del inglés (Çelik, Aytin & Bayman, 2013; Ning, 2013; Tuan, 2013; Wang & Hu, 2009), al aprendizaje en grupos cuando se ha aplicado a metodologías de enseñanza a distancia usando bandas estrechas de Internet (Tilwaldi, Takahashi, Takata & Koizumi, 2010) o para la integración de estudiantes con discapacidades auditivas al usarse Internet como herramienta para el desarrollo de la enseñanza con métodos cooperativos (Sheng & Xu, 2009).

Son otros ejemplos, el uso de blogs en la enseñanza en Educación Superior para buscar efectos favorables en interacciones sociales, tiempo de respuesta y cantidad de conocimiento adquirido (Huang, Huang & Yu, 2011); las plataformas de enseñanza virtual que favorecen una comunicación interpersonal cooperativa (Hassan, Fong & Idrus, 2011; Roseth, Saltarelli & Glass, 2011); el uso de herramientas tecnológicas como apoyo a los métodos de aprendizaje cooperativo para el desarrollo de competencias en la enseñanza a distancia en Enfermería (Lin, 2012); el diseño de informes de actividades on-line cooperativas (Lo, 2013) y la continua revisión del método *Computer-Supported Collaborative Learning* (CSCL) para potenciar el rendimiento (Resta & Laferrière, 2007).

Inclusive, apunta Millis (2010b) que cuando en las facultades se introducen elementos como las nuevas tecnologías éstas se usan ampliamente porque incrementan el tamaño de las clases gracias a la virtualidad, lo que permite la asistencia a estudiantes de lugares diferentes favoreciendo así la cooperación.

El trabajo por parejas también resulta interesante en la universidad. El método *Reciprocal Peer Tutoring* mejora la metacognición (De Backer, Van Keer & Valcke, 2012) porque proporciona una ganancia importante en los niveles de motivación que justifica los costes de aplicación (Rittschof & Griffin, 2001) y el impacto del método *Peer Tutoring* incide en el rendimiento, las estrategias de aprendizaje y en las mejoras en torno a la relación con los tutores y las estrategias de aprendizaje (Young, 2011).

#### *Aprendizaje cooperativo entre estudiantes, profesores y Facultad*

Autores dedicados durante muchos años a la vida universitaria en la comprensión de los factores que inciden en el aprendizaje en Educación Superior, la calidad e innovación educativa, las causas y efectos del abandono escolar y en especial, la relación del

estudiante no solo con su grupo de clase y el profesor o los profesores sino con la Facultad y a la vez, con otros estudiantes de otras Facultades de otras universidades, confirman la necesidad del trabajo cooperativo, en especial, en los primeros cursos académicos (Goodman & Pascarella, 2006; Pascarella et al, 2007; Tinto, 1997).

El trabajo clásico de Tinto (1993) identifica ocho causas (una involuntaria y siete voluntarias) que explican las razones por las cuales los estudiantes abandonan la universidad antes de graduarse: bajo rendimiento académico, mala adaptación a la vida universitaria, incertidumbre acerca de las metas personales, falta de compromiso con la universidad, dificultades económicas, no integración social en la vida universitaria, incongruencia entre los planes de estudios universitarios y los intereses de los estudiantes y el aislamiento social e intelectual.

Añade que en la medida en que se utilice el aprendizaje cooperativo con mayor frecuencia en la universidad, ésta será mas exitosa. El rendimiento académico promovido por la cooperación puede aumentar el éxito académico de los estudiantes, lo que reduce el fracaso escolar.

Las relaciones interpersonales positivas impulsadas por experiencias cooperativas pueden disparar la presión social para aprender, adaptarse a las nuevas relaciones y a la inserción social en la vida del campus, a establecer objetivos sociales, reducir la incertidumbre acerca de las metas y a aumentar el compromiso con otros estudiantes. El incremento de la salud psicológica promovida por las experiencias de cooperación académica puede elevar la autoestima de los alumnos y la auto-eficacia, clarificar las metas personales, acrecentar la capacidad de hacer frente a la incertidumbre, mantener relaciones constructivas con diversos compañeros, formar coaliciones para alcanzar los objetivos y desarrollar la capacidad para adaptar los objetivos personales a las situaciones actuales (Tinto, 1993).

Por tanto, entre los factores que caracterizan las buenas prácticas universitarias, el aprendizaje cooperativo ocupa el sexto lugar, después de (a) la enseñanza e interacciones de alta calidad con el profesorado, (b) diversidad de experiencias, (c) el desarrollo académico y las altas expectativas, (d) la frecuencia de las interacciones con los profesores/profesionales y (e) las interacciones con sus compañeros (Cruce, Wolniak, Seifert & Pascarella, 2006; Pascarella, Wolniak, Seifert, Cruce & Blach, 2005; Pascarella et al. 2006).

Los 4501 estudiantes del primer curso de 19 universidades americanas que participaron en el estudio de Seifert, Goodman, King & Baxter (2010) confirman la interacción entre pares como un factor clave para la transición a la universidad, resultado que también obtuvo Tinto (1993) en sus hallazgos: “the *process* of collaborative learning is as important as is content” (p. 168) a lo que agrega: “the evidence is clear that collaborative/cooperative learning settings are more likely than traditional lecture settings to engage the energies of adult students” (p. 189). Los alumnos de nuevo ingreso que han trabajado cooperativamente en el aula incrementan el tiempo fuera de clase para trabajar con sus pares.

Avalando las evidencias afectivas, sociales y cognitivas de esta metodología, el estudio reciente de Svinicki & McKeachie (2014) ofrece siete sugerencias a los estudiantes universitarios para lograr grupos efectivos: (a) cada miembro contribuye a la discusión y asume su tarea dentro del grupo, (b) cada integrante tiene que asegurarse que todas las ideas, por muy pequeñas y sencillas han sido consideradas, (c) no asumir que hay consenso porque no hay ideas opuestas o alternativas, (d) determinar objetivos a corto, mediano y largo plazo pero sin tener miedo al cambio durante el progreso, (e) estar seguros de que cada persona conoce su tarea y la de los otros y el tiempo límite de entrega (f) acordar entre todos el día y hora para el próximo encuentro y (g) antes de terminar la reunión hay que evaluar el progreso del grupo y qué otras estrategias se pudieran hacer de manera diferente para el próximo encuentro.

No obstante, la experiencia de Cohen (2010) quien ha dictado asignaturas bajo criterios de aprendizaje cooperativo a lo largo de 25 años en tres universidades norteamericanas, sale del aula para invitar al profesorado a una reflexión permanente sobre su propia práctica, proposición que es apoyada por Park & Choi (2014) quienes afirman de manera desmesurada pero sin dejar de tener algo de razón

Educational environment influences students' learning attitudes, and the classroom conveys the educational philosophy. The traditional college classroom design is based on the educational space that first appeared in medieval universities. Since then classrooms have not changed except in their size. In an attempt to develop a different perspective of educational environment, a new design of classroom, the active learning classroom (ALC) (p.749).

Enunciado que amplía Plank (2011) quien recomienda el trabajo cooperativo entre profesores y la creación de líneas de investigación, a partir de las reflexiones de los

profesionales, que hacen a las asignaturas espacios de cuestionamiento para el aprendizaje.

Las experiencias a dúo, tan enriquecedoras como las de profesores que en solitario desarrollan su labor, son para el aprendizaje cooperativo mecanismos de comprensión que no sólo trabajan en grupo los alumnos, sino que bien articuladas, el trabajo en pares entre profesionales de la docencia genera frutos oportunos y, por demás, enriquecedores para lo que, en materia de conocimiento de la enseñanza universitaria se trata (Jessen–Marshall & Lescinsky, 2011). Tan es así, que el estudio de Clarke, Trigss & Nielsen (2014) confirma que “Student teachers consider cooperating teachers to be one of the most important contributors to their teacher preparation program. Therefore, the ways in which cooperating teachers participate in teacher education are significant” (p. 163).

Es cierto que la puesta en práctica de esta metodología requiere superar las dificultades propias de un método que no es tan habitual en el ámbito universitario, lo que puede provocar reacciones de hostilidad entre los estudiantes. La ausencia constante o la pérdida de un miembro del grupo durante el semestre, el no contar con un aula que favorezca las interacciones cara a cara, o el no seguimiento minucioso de las interacciones entre los grupos cooperativos de base en donde se trabaja por periodos largos de tiempo, pueden mermar la cohesión grupal o la efectividad de los mismos (Bryant, 1978).

As with all kinds of teaching, designing and guiding group work takes time to learn and practice. And for students, learning to learn well in groups doesn't happen overnight. Most teachers start with modest efforts. Many work with colleagues, designing, trying and observing each other's approaches (Smith & MacGregor, 1992, p. 26).

Por consiguiente, Huber (2008) sugiere cuatro principios a seguir para evitar las dificultades de la puesta en marcha de métodos cooperativos: ayudar a los alumnos con estrategias adecuadas de aprendizaje, promover la enseñanza mutua, retroalimentar el rendimiento del grupo y no de los miembros individuales e iniciar procesos de reflexión con los estudiantes

Pero la meta general del aprendizaje activo, situado, autorregulado, constructivo y social confronta tanto a estudiantes como a profesores con la dificultad común que estas mismas características determinan al mismo tiempo condiciones previas, que el principiante debería traer en la situación del aprendizaje activo, autorregulado, etc. Es decir, cuando un profesor prepara situaciones de aprendizaje activo para sus estudiantes, que no disponen ya de destrezas de aprender activamente, meterá a los estudiantes en líos. Por eso, sobre todo los estudiantes acostumbrados a recibir la información preparada para poder

asimilarla necesitan estar dirigidos en sus actividades y reflexionar sobre su progreso y sus problemas (Huber, 2008, pp. 71–72).

Los métodos activos que son adecuados para ser trabajados en el Espacio Europeo de Educación Superior son aquellos que preparan en conocimiento y en competencias a los estudiantes para el mundo laboral, a ser ciudadanos activos en una sociedad democrática y para que asuman responsabilidades individuales (Huber, 2008). Sin duda, el aprendizaje cooperativo es uno de los métodos que cumple con estas tres condiciones.

### **2.1.2. Perspectiva española de investigaciones en la universidad**

Centrándonos en España, hay acercamientos en la década de los años 70 y 80, como lo es el caso de García Hoz (1972), Medina Rivilla (1979), el movimiento de renovación pedagógica de la conocida “Experiencia Somosaguas” (Sánchez, 1996 y Gutiérrez, en el año 2009 retoma la experiencia), Coll (1984) y Villalvilla, López y Llamas (1988) o el libro de Arnaiz sobre el aprendizaje en grupo (1989).

Pero realmente el aprendizaje cooperativo comenzó a ser estudiado en tierra ibérica en los años 90, dos décadas más tarde si se compara con el inicio en los Estados Unidos de Norteamérica con autores como Ovejero (1990a, 1990b, 1993), Colomina y Onrubia (1990), Echeita y Martín (1990), Echeita (1995), Santos Rego (1990), Fabra (1994), Díaz–Aguado (1996), Serrano y González–Herrero (1996) y Lobato (1998). Para aquel momento, estos precursores ya hacían mención de la necesidad de la implementación de esta metodología en las aulas españolas y algunos de ellos específicamente reclamaban su puesta en práctica en el ámbito universitario.

Las investigaciones de ese entonces se agruparon en dos monográficos dedicados al tema en donde se evidencia la preocupación española por formar parte del debate en innovación educativa (Barnett, 1995 citado por Cassany, 2004). La revista *Aula de Innovación Educativa*, en el año 1992, con el título “El trabajo en grupo” y la revista *Cuadernos de Pedagogía*, en el año 1997, con los rótulos de “Colaborar para aprender” y “Cooperación y diversidad”.

Esta última década del siglo pasado refleja un volumen de investigaciones cooperativas en los anaqueles de las bibliotecas españolas sobre la fundamentación teórica en aprendizaje cooperativo (Lobato, 1997; Martí, 1997) y su eficacia (Martí, 1992; Martí 1997; Martí y Solé, 1997); los recursos prácticos cooperativos (Pallarés, 1990); la integración social, especialmente niños de raza gitana y niños discapacitados (Cambra y

Laborda, 1998; Díaz–Aguado y Andrés, 1999; Ovejero, Gutiérrez y Fernández Alonso, 1996); el estudio de variables asociadas como el desarrollo moral (Ortega, Mínguez y Gil, 1997), el autoconcepto (Uriarte, 1996), las habilidades de comunicación (Solsona, 1999) y el rendimiento académico (Santos Rego, 1999) o la apuesta por la innovación educativa en la enseñanza de las matemáticas (Gavilán, 1997; Serrano, González–Herrero y Martínez–Artero, 1997) y el aprendizaje de la Música (Llopis, 1999; Puchau, 1999).

En la primera década del siglo XXI se han seguido sumando otras voces consolidando así el esfuerzo inicial en el estudio de los conceptos básicos del aprendizaje cooperativo (Echeita, 2012; León del Barco, Gozalo, Felipe, Gómez y Latas, 2005; Negro, Torrego y Zariquiey, 2012) y su implementación en las aulas (Duran, 2012; Más, Negro y Torrego, 2012; Pujolàs, 2009, 2012); las habilidades sociales y los factores que inciden en la interacción (León del Barco, 2006; León del Barco, Gozalo y Vicente, 2004; Pérez, 2003); la atención a la diversidad (López, Ojea y Fernández, 2000; Pujolàs, 2004); métodos específicos para la enseñanza de las matemáticas (Gavilán, 2004, 2009; Gavilán y Alario, 2010; Pons, González–Herrero y Serrano, 2008; Serrano, González–Herrero y Pons, 2008); la tutoría entre iguales (Duran, 2006; Duran, Torró y Vila, 2003; Duran y Vidal, 2004; Sáiz y Román, 2011); la cooperación mediada por las TIC (Adell, 2011; González y Zariquiey, 2012; Marqués, 2011) o las diferencias entre aprendizaje cooperativo y colaborativo (Duran y Monereo, 2012).

Pese a que autores como Goikoetxea y Pascual (2002) y León del Barco, Gonzalo y Vicente (2004), consideran que son limitados los estudios sobre aprendizaje cooperativo en España y que los contados hallazgos no han buscado identificar los factores que median su eficacia, hay que agregar a esta acotación que la mayoría de esa producción responde a investigaciones descriptivas y no a estudios experimentales.

Así, versan trabajos descriptivos sobre las características del estudiante y su valoración frente al aprendizaje cooperativo (Callado y Utrero, 2012; López, 2012; Muela, 2012; Sánchez y Pueo, 2012); relativos a la interacción del alumnado on–line (Hernando, Aguaded y Tirado, 2011; Vaquero y de Lorenzo, 2009); la valoración de los estudiantes con respecto al método aprendizaje colaborativo asistido por ordenador (CSCL) (García y Suárez, 2011; Gómez, Palomares y Pino, 2010; Salmerón, Rodríguez y Gutiérrez, 2010; Serrano, Muñoz y López, 2012); experiencias docentes para la mejora de la enseñanza y buenas prácticas (Gutiérrez, Yuste, Cubo y Fustes, 2011; Reyes y Gálvez,



2010); la comunicación sincrónica gracias a las TIC (Trigueros, Rivera y De la Torre, 2011); el contraste entre el rendimiento de alumnos que se forman de modo virtual frente al presencial con el uso de metodologías en aprendizaje cooperativo (Rivero, Samino y Pérez del Campo, 2008); la evaluación cooperativa (Jiménez, 2012; López y Álvarez, 2011; Palomares, 2011) o la coevaluación (Gessa, 2011).

En materia de otros métodos propios del aprendizaje cooperativo, el *Peer Tutoring* ha ofrecido interesantes resultados a favor del tratamiento educativo en grupos porque mejora el rendimiento y las estrategias de aprendizaje (Arco-Tirado, Fernández-Martín & Fernández-Balboa, 2011). Los ingenieros son otro ejemplo de cómo en sus clases se aplican métodos cooperativos para la enseñanza (Gil, Baños, Montoya, Herrada y Montoya, 2013; Vergara, 2012) siendo una metodología que se ha aplicado no solo en aulas convencionales sino en proyectos reales Aeroespaciales en España (Rodríguez, Miaja, Arias, Rodríguez y Lamar, 2010) o en las Ciencias de la Salud como lo es la Enfermería (Molina, Gil, Bravo y Herrero, 2010; Ulibarri y Fernández, 2010).

La formación del profesorado universitario y los grupos cooperativos docentes son otra vertiente de gran interés (Boronat y Ruiz, 2011; Escudero, 2012; Fraile, 2008; León del Barco, Felipe, Iglesias y Latas, 2011; León del Barco y Latas, 2007; Rayón, 2012; Iborra y Dasí, 2009; Sanz y Reig-Hernández, 2013).

En esta materia, un estudio bibliométrico da cuenta de la producción realizada entre 1997 y 2008 en la formación del profesorado en la implantación de técnicas de aprendizaje cooperativo en el aula (Pérez, Paz y Poveda, 2009); otro destaca la formación continua en métodos de aprendizaje cooperativo entre universidades españolas y del extranjero (Cordero y Serrano, 2010); otro la participación de varias facultades universitarias en proyectos cooperativos de aprendizaje (Lobato, Apodaca, Barandiaran, San José, Sancho y Zubimendi, 2010) y hay quienes estudian la cooperación como una competencia a desarrollar (Apodaca, 2009; Benito, Bonson e Icarán, 2005; López, 2005; Medina Rivilla, 2009; Medina Rivilla, Sevillano y de la Torre, 2009; Zabalza, 2004; 2007). Se suman bajo este mismo interés estudios que aluden a propuestas de renovación metodológica en pequeños grupos en donde la función tutorial, los conceptos de formación competencial y el aprendizaje autónomo son fundamentales (Álvarez y García, 2012; Díez, 2011; Andreu, Sanz y Serrat, 2009).

Sin embargo, hay que ser conscientes de que el trabajo en equipo viene acompañado de las dificultades propias de su puesta en marcha en el contexto universitario, realidad que tampoco escapa al escenario español. La masificación de alumnos, aulas no condicionadas con un mobiliario apropiado para promover un ambiente cooperativo, incorporación tardía de estudiantes a los grupos, inasistencia a clase, baja participación en el trabajo personal y una mayor carga para el profesorado (De Arriba y García, 2013; Siota, 2009) son obstáculos que pueden ser asumidos por profesores y estudiantes como un reto permanente (Boza y Toscano, 2012)

Los alumnos no están acostumbrados a trabajar juntos y las tareas de aprendizaje cooperativo no les resultan cómodas y les requiere una mayor dedicación. Tampoco los profesores estamos habituados a este tipo de enseñanza y su implementación supone un gran esfuerzo sobre todo para grupos docentes con un número tan elevado de estudiantes (Molina, Gil, Bravo y Herrero, 2010, p. 267).

Es necesario que la universidad española, mucho mas orientada en valorar los logros en investigación aunque diseñada para alcanzar resultados pedagógicos (Sancho, 2001), coloque al mismo nivel que sus éxitos científicos la excelencia en la enseñanza (Bolívar y Caballero, 2008), en donde la formación continua en innovación educativa cumpla un papel prioritario porque

el profesorado universitario tiene amplia experiencia en cuanto a cooperación en el desarrollo de proyectos de investigación. Sin embargo, no ocurre lo mismo con la cooperación para el desarrollo de la actividad docente, lo que dificulta la renovación de las actuales metodologías (Ministerio de Educación y Ciencia, 2006, p. 48)

Por tanto, el listado no exhaustivo de investigaciones aquí presentadas es un panorama que refleja el quehacer del profesor universitario en su función docente–investigadora en donde:

- (a) es necesario que la universidad apoye la formación del profesorado para guiar las situaciones de aprendizaje cooperativo siendo un requisito acompañar esta preparación con un apoyo institucional que decante en recursos físicos y económicos (León del Barco y Latas, 2007; Santos Rego, 1990),
- (b) son escasos los estudios que midan la eficacia del aprendizaje cooperativo en las aulas universitarias españolas (León del Barco y Latas, 2007) y
- (c) cada vez mas la puesta en práctica de esta metodología en España demuestra el interés por su difusión en revistas científicas internacionales y por ende, en el Espacio

Europeo de Educación Superior para así dar a conocer lo que se está haciendo desde la universidad en materia de innovación educativa (Arco-Tirado, Fernández-Martín & Fernández-Balboa, 2011; Alorda, Suenaga y Pons, 2011; Coll, Rochera y Gispert, 2014; Fernández, 2004; León del Barco, Felipe, Iglesias y Marugán, 2014; Lobato et al, 2010; Parra-Meroño y Peña-Acuña, 2012; Rodríguez, Miaja, Arias, Rodríguez y Lamar, 2010).

## **2.2. Eficacia del aprendizaje cooperativo en el rendimiento: meta-análisis**

En los años 1940 y 1950, en las grandes publicaciones como *Review of Educational Research* y *Psychological Bulletin*, se podían encontrar gran cantidad de investigaciones sobre un mismo tema. Para aquel momento, la integración teórica de pocos estudios a través de revisiones narrativas fue satisfactorio, a pesar de las limitaciones propias, debido a la falta de sistematización en la localización de los estudios, la no delimitación clara de los objetivos, la falta de evaluación de la calidad de las investigaciones originales, la carencia de síntesis de los datos cuantitativos y la ausencia de una interpretación objetiva de los resultados encontrados (Glass, 1976, 1977).

Años después, en 1960, la investigación creció exponencialmente. Los investigadores continuaron integrando los estudios de manera narrativa pero estas descripciones cronológicas verbales ya no estaban representando el conocimiento acumulado. En consecuencia, los investigadores comenzaron a hacer clasificaciones y mediciones de las condiciones y los resultados de los estudios. Con frecuencia, fueron clasificadas en tablas por tipos y en función de resultados estadísticamente significativos.

La integración de la investigación en la literatura, en los años 1970, demandaba mayor sofisticación en las técnicas de medición y análisis estadístico que fueran más allá de la retórica y la narrativa

I had hope to find research to support or to conclusively oppose my belief that quality integrated education is the most promising approach. But I have found very little conclusive evidence, for every study, statistical or theoretical, that contains a proposed solution or recommendation, there is always another, equally well documented, challenging the assumptions or conclusions of the first. No one seems to agree with anyone else's approach. But more distressing: no one seems to know what works. As a result I must confess, I stand with my colleagues confused and often disheartened (Mondale, 1970 citado por Hunter & Schmidt, 1996 en 78th Annual Convention of the American Psychological Association).

Vista la necesidad, Glass (1976, 1977) creó un enfoque que se distinguiera de los que hasta el momento se habían desarrollado. Decidió llamar a este enfoque meta-análisis de la investigación.

En palabras del propio autor, se trata de un término polémico pero a la vez preciso y pertinente el cual debe diferenciarse –vista la rigurosidad que lo caracteriza– del análisis narrativo, tan característico de los intentos por ofrecer un sentido a la literatura científica o de otros como la meta-etnografía (*meta-ethnography*) o el conteo de revisiones (*vote counting reviews*). Dos décadas mas tarde y en la actualidad, la sofisticación del análisis estadístico y del conjunto de datos a partir de los estudios primarios ha sido la tendencia (Cook et al. 1992; Cooper & Hedges, 1994; Kulik & Kulik, 1989; Lipsey & Wilson, 2001).

Así el meta-análisis se define como el análisis estadístico de una larga colección de resultados, relativos a un problema de investigación, provenientes de estudios individuales con el propósito de integrar sus conclusiones. Coloquialmente se le conoce como el análisis de los análisis y como cualquier otro método no escapa a sus propias limitaciones: la especificidad de la cuestión que se está dirigiendo y la amplitud de la literatura y de datos de búsquedas en los que se basa (Davies, 2000).

Pero ¿cuándo el meta-análisis no es la metodología apropiada? Botella y Gambara (2006) lo resumen en tres circunstancias: (a) los estudios que se pretenden integrar provienen de una perspectiva cualitativa, (b) se intenta ofrecer una respuesta a la dirección causal de la relación entre dos variables y (c) las investigaciones primarias no son suficientemente homogéneas –aunque no existe un criterio para establecer ese nivel “idóneo” de homogeneidad– (Cooper, 2003).

En la universidad, los estudios meta-analíticos siguen constituyendo una pequeña proporción de la revisión de la literatura y los pocos que existen no han sido publicados en revistas o libros relacionados con la Educación Superior. Quizás, la falta de conocimiento acerca de cómo llevar a cabo un meta-análisis sea una de las razones principales de esta ausencia y la otra, el reto al que se enfrentan los académicos universitarios porque “higher education scholars face challenges regarding effect sizes and independence of observations that are often less problematic in other fields of study” (Bowman, 2012, p. 376).

Sin embargo, la razón de ser del meta-análisis es que trascienda los resultados obtenidos para la toma de decisiones porque

The results of meta-analytic research can be quite helpful not only to researchers who want to make sense of disparate findings that sometimes occur across studies, but also to administrators and practitioners who are interested in improving their institutions and the experiences of their students and faculty (Bowman, 2012, p. 380).

### **2.2.1. Evidencia Basada en Investigación Educativa**

En Educación, a diferencia de otras áreas de conocimiento, se han tomado decisiones en muchas ocasiones carentes de evidencia y sustento empírico. Por lo general, el criterio que respalda la orientación de los programas de instrucción, los libros de texto, las metodologías didácticas, la formación del profesorado y las reformas educativas que luego se concretan en el quehacer del docente en el aula, han sido objeto de la ideología, la política, el *marketing* y la moda. En fin, de prácticas que se generalizan a pesar de no contar con un respaldo investigativo que repercuta en una educación de calidad.

Precisamente, esa toma de decisión debe sustentarse en la evaluación de síntesis de investigaciones que demuestran a partir de la revisión de numerosos estudios primarios la eficacia y eficiencia del programa educativo que se desea valorar, valoración que a su vez se soporta en criterios que garanticen una información lo más confiable, imparcial y significativa posible. Considerar el diseño de la investigación, el tamaño de la muestra, el ajuste de los grupos antes de la aplicación de la prueba, la duración de la intervención y las medidas aplicadas, son cuestiones metodológicas y sustantivas que no difieren de otras síntesis cuantitativas, pero que a su vez, cuando se llevan a la evaluación de programas educativos deben adaptarse a las características propias de ese campo de estudio (Slavin, 2008).

Al comparar la investigación educativa con otros marcos de referencia Slavin y Fashola (1998) se cuestionan “why is it that medicine, engineering, agriculture –in fact, most fields of endeavor– make steady progress over time in their basic technologies and effectiveness, whereas education moves from fad to fad with little apparent impact on student outcomes?” (p. 6). A su vez, Hargreaves (2011) se pregunta por qué si la Educación y la Medicina son profesiones centradas en la persona, si ambas están comprometidas con la calidad y el progreso de las naciones, la Medicina ha alcanzado un reconocimiento y prestigio social y académico mientras que la Educación ha quedado rezagada en sus propuestas educativas.

Para los primeros autores, Slavin y Fashola (1998) “the most important reason is that in other fields, research is respected and used as a guide to practice, whereas in education this is hardly the case” (p. 6). Para el segundo, Hargreaves (2011), razones de peso acercan y alejan a las dos profesiones desde un punto de vista investigativo pero las une un hilo común: la mejora de la práctica en beneficio del otro.

En el siglo XXI, la evidencia como base para la adopción de programas y prácticas educativas, llamada *Evidencia Basada en Investigación Educativa* –en inglés *Evidence-Based Research in Education*– comienza a promoverse cada vez con mayor fuerza. Educadores, padres, estudiantes e investigadores comparten una misma interrogante ¿en qué medida las prácticas educativas que actualmente se utilizan en las escuelas son efectivas y eficientes? Se habla entonces de la investigación al servicio de las buenas prácticas, de los estudios que utilizan métodos rigurosos y diseños experimentales aleatorizados para orientar una política educativa de calidad.

La Evidencia Basada en Investigación Educativa se entiende como la integración de los conocimientos profesionales con la mejor evidencia empírica disponible para la toma de decisiones acerca de cómo impartir la enseñanza “the integration of professional wisdom with the best available empirical evidence in making decisions about how to deliver instruction” (Moran, 2004, p. 4). Requiere tanto de la “evidencia empírica” como de la “sabiduría del profesional”. Lo primero, responde a cuáles de los métodos de enseñanza son fiables de llevar a cabo para la mejora educativa y a su vez, cuáles de ellos pueden implementarse en un corto tiempo o con la menor disponibilidad de recursos. Lo segundo, atiende a cómo el profesor es capaz de adaptar apropiadamente esos métodos a las características propias del contexto educativo en donde se desenvuelve con su grupo de estudiantes. Es más, la profesionalidad del docente viene dada por su capacidad de incorporar esas metodologías a un currículo, a un plan de estudios de un determinado entorno.

La mayor trayectoria en meta-análisis y revisiones sistemáticas proviene del campo de la Medicina y de la Salud (Davies, 2000) a través de redes internacionales de investigación como *The Campbell Collaboration* y *The Cochrane Collaboration* pero también es cierto, que existe una tradición de grupos de investigación y organizaciones fundamentalmente norteamericanas en Educación que han trabajado desde hace varios años en la elaboración de indicadores de calidad para la valoración de estudios sustentados en evidencias basadas en la investigación como son *What Works*

*Clearinghouse* (WWC) del Institute of Education Science, *Best Evidence Encyclopedia* (BEE) de Johns Hopkins University y *The Center for Comprehensive School Reform and Improvement* del U.S. Department Education y (para ver los criterios de calidad que utilizan estas organizaciones ir al anexo 2).

En Europa, también existen iniciativas como *The Evidence for Policy and Practice Information and Co-ordinating Centre* (EPPI-Centre) del Instituto de Educación de la Universidad de Londres y el Proyecto de Políticas y Prácticas en Educación Basadas en la Evidencia en Europa (*Evidence Informed Policy and Practice in Education in Europe*, 2011 –EIPPEE, por sus siglas en inglés–), con la participación de 23 países europeos y siete organizaciones con sede en cuatro países no europeos, interesados en estudiar la naturaleza y el alcance de actividades que vinculan la evidencia de investigación con la formulación de políticas educativas en la comunidad europea.

En España, los marcos de referencia se sitúan por su tradición en la Medicina Basada en la Evidencia (EBM) y recientemente, en la Psicología Basada en la Evidencia (Llobell, Navarro y Monterde, 2004; Navarro y Llobell, 2003 ver estos autores en Sánchez-Meca y Botella, 2010), la Intervención Psicosocial Basada en la Evidencia (Sánchez-Meca, Marín-Martínez y López-López, 2011), la Práctica Basada en Evidencias (Grupo de Atención Sanitaria Basada en la Evidencia, 2007; Navarro, Giribet y Aguinaga, 1999; Vázquez y Nieto, 2003 ver estos autores en Sánchez-Meca y Botella, 2010) y la Salud Mental Basada en la Evidencia (EBMH). Es un panorama que reclama la consolidación de una Evidencia Basada en Investigación Educativa y, a la vez, el emprender una investigación fundada en evidencias para la enseñanza universitaria centrada en la innovación docente (Becerra et al. 2012).

Entonces si “... a whole industry of systematic reviews and meta-analyses has developed in educational research, particularly in the USA. The question to be asked is: what has been learned from all this industry of systematic reviews and meta-analyses in education?” (Davies, 2000, p. 372). Defensores y detractores de esta metodología de investigación matizan sus bondades y limitaciones.

Kulik y Kulik (1989) y Lipsey y Wilson (1993) quienes han revisado las mejores evidencias en Educación –150 meta-análisis relacionados con intervenciones en Educación y 302 meta-análisis en intervenciones en Psicología y Educación, respectivamente– son los primeros en cuestionar los resultados encontrados.

Afirman que la magnitud del efecto de esas intervenciones en su mayoría son moderados y los efectos fuertes o negativos no ocurren con frecuencia, lo que hace preguntarse a estos autores si se trata de un artefacto de la metodología de las revisiones sistemáticas, de los tipos de diseños de investigación, si refiere a la calidad de los estudios primarios, o a sesgos en el tamaño de la muestra o de publicación. Otros temas relacionados como la significación estadística y la significación educativa, la pertinencia de algunas medidas de resultados del logro académico que presentan dudas en cuanto a su validez ecológica externa e interna y la generalización de los resultados, plantean dificultades que inciden directamente en la toma de decisiones para una práctica educativa de calidad (Davies, 2000).

Para Slavin (1984, 1986) no considerar como criterio a priori solo la selección de aquellos estudios primarios con una alta calidad metodológica es un signo de debilidad. La premisa *best evidence syntheses* propia de este autor y que acompaña reiterativamente sus revisiones narrativas y meta-análisis resulta un tema recurrente de discusión porque esta regla hace que se abarquen contados estudios y la selección de estos no esté exenta de prejuicios del analista (Kulik & Kulik, 1989).

Por su parte, Clegg (2005) aboga para la superación de estas dificultades, una perspectiva crítica realista de la práctica basada en la evidencia en donde “...thinking from a critical realist perspective liberates the space for theoretically informed work, whereby arguments about method, and in particular randomized controlled trials, do not become a proxy for the open examination of ontological and epistemological assumptions” (p. 415) argumento que deja entrever la necesidad de seguir trabajando el meta-análisis como un método de investigación que mejora la práctica educativa

these problems exist also for the narrative review, and that the key advantage of the systematic approach of a meta-analysis is that all steps are clearly described so that the process is transparent (...) Most of the criticisms raised deal with the application of the method, rather than with the method itself. What we should do is take the valid criticisms seriously and protect against them in planned analyses and by thoughtful interpretation of results (Borenstein, Hedges, Higgins & Rothstein, 2009, p. 386).

Pero “¿What would happen to education if we take education evidence seriously?” Ocorre que trasciende el punto de vista de la evidencia empírica derivada de la investigación educativa y de otros ámbitos relacionados porque incluye las teorías en que se sustenta esa evidencia empírica. Son las teorías las que ofrecen explicaciones del



por qué el proceso de enseñanza–aprendizaje puede funcionar en Educación (Van der Vleuten & Driessen, 2014).

Precisamente, es la Evidencia Basada en Investigación Educativa la que confirma que el aprendizaje cooperativo es un área educativa con evidencias sólidas porque: (a) resulta mas efectivo que el individual y (b) se conocen ciertas condiciones bajo las cuales el aprendizaje es mas productivo (Van der Vleuten & Driessen, 2014), testimonio que impulsa a demostrar a partir de una revisión cuantitativa el efecto del aprendizaje cooperativo focalizado en el rendimiento de estudiantes universitarios porque “...despite the volume of research on cooperative learning, few investigations have focused on college students outside the psychology laboratory. To our knowledge, no meta–analysis of small–group learning focuses exclusively on undergraduates” (Springer, Stanne & Donovan, 1999, p. 3).

### **2.2.2. Meta–análisis en los que se incluye a la universidad**

Existe una larga tradición de meta–análisis de situaciones competitivas, cooperativas e individuales, empezando por los años 1980 por Turner en Reino Unido y Triplett en los Estados Unidos y en Alemania, en los años 1990 por Mayer y en Francia por Ringelmann (Johnson & Johnson, 2002a) aunque el mayor volumen de estudios primarios en las revisiones cuantitativas se concentre en el aprendizaje cooperativo versus el aprendizaje individual, por tratarse este último del aprendizaje que ha prevalecido en las aulas universitarias

With the amount of evidence available, it is surprising that practice in college classrooms is so oriented toward competitive and individualistic learning and college are so dominated by competitive and individualistic organizational structures. It is time for the discrepancy to be reduce between what research indicates is effective in teaching and what college faculty actually do. To do so, faculty must understand the role of the instructor in implementing cooperative learning (Johnson, Johnson & Smith, 1991, p. 55).

Retomando parte de la aseveración alegada por Springer, Stanne & Donovan (1999) “To our knowledge, no meta–analysis of small–group learning focuses exclusively on undergraduates” (p. 3), es cierto que son contados los meta–análisis en aprendizaje cooperativo que se han dirigido exclusivamente al ámbito universitario. Una muestra de esas escasas revisiones cuantitativas se describe en las siguientes líneas.

Un meta–análisis clásico es el de Johnson & Johnson (1989). Su libro *Cooperation and Competition: Theory and Research* es en sí mismo una síntesis de los resultados

obtenidos a partir de la revisión de 529 documentos desde el año 1898 hasta el año 1989. Compararon en todos los niveles educativos el aprendizaje cooperativo con situaciones competitivas e individuales con respecto al rendimiento académico, relaciones interpersonales, apoyo social, bienestar psicológico y autoestima.

En cuanto al rendimiento, obtuvieron una magnitud del efecto global a favor de la cooperación versus situaciones competitivas ( $d=0,67$ ,  $n=185$ ) e individuales ( $d=0,64$ ,  $n=226$ ). Cuando el análisis se reduce solo a aquellos estudios con alta calidad metodológica la evidencia es mas fuerte a favor del aprendizaje cooperativo en comparación con el competitivo ( $d=0,88$ ,  $n=51$ ) mientras que la magnitud del tamaño del efecto se mantiene similar con respecto al aprendizaje individual ( $d=0,61$ ,  $n=104$ ).

No siendo el objetivo principal el análisis por etapas educativas, son dos las alusiones a las que hacen referencia estos autores para el nivel de Educación Superior. Una de ellas confirma que en los 137 estudios experimentales que se llevaron a cabo en la universidad y con adultos, el aprendizaje cooperativo favorece el rendimiento en contraparte con la competitividad y la individualidad ( $d=0,59$  y  $0,62$ , respectivamente). La otra, la dedican a los estudios cuyos resultados son opuestos a favor del aprendizaje cooperativo. Por ejemplo, que solo en ocho estudios la competición obtuvo un nivel de rendimiento mayor que la cooperación y que en cinco estudios el rendimiento en el aprendizaje individual prevaleció.

Los autores dejan muy claro en sus conclusiones que estos hallazgos negativos obedecen al nivel educativo y al tipo de tarea desempeñada “all of the studies used college and adult samples and the tasks were either brainstorming or simple motor and rote tasks” (Johnson & Johnson, 1989, p. 171) y no a la eficacia del aprendizaje cooperativo, el cual ha sido confirmado en numerosas ocasiones porque promueve los procesos más altos de razonamiento y la transferencia de conocimiento.

Johnson, Maruyama, Johnson, Nelson y Skon (1981) analizaron 122 estudios para comparar la efectividad de estructuras de aprendizaje centradas en la cooperación, la cooperación con grupos intercompetitivos, la competición interpersonal entre los miembros del grupo y el esfuerzo individual. Los hallazgos confirmaron que la cooperación es superior cuando se compara con situaciones individuales en estudiantes de primaria, secundaria, universidad y adultos ( $d=0,78$ ,  $DE=0,91$ ,  $n=104$ ). Si bien es cierto la investigación menciona solamente las características generales de los estudios

primarios, no presenta el porcentaje de documentos seleccionados por nivel educativo, ni tampoco un análisis por grupos que permitiera conocer lo que ocurrió específicamente en el ámbito universitario.

Otro estudio dirigido por Qin, Johnson y Johnson (1995) quienes no contrastaron situaciones cooperativas con situaciones individuales pero si con la competición ante la resolución de problemas, ofreció resultados interesantes en cuanto a las tareas diseñadas.

Estos autores analizaron 63 resultados de 46 estudios entre 1929 y 1993, de los cuales solo 33 resultados se llevaron a cabo entre secundaria, universidad y adultos. Los documentos primarios fueron clasificados en cuatro categorías de acuerdo al tipo de tarea a resolver: lingüística (para ser respondidas a través del lenguaje oral y escrito), no lingüística (para ser respondidas a través de símbolos, matemáticas, actividades motoras y acciones), estructurada (aquellas que poseen una definición clara de las operaciones y soluciones) y no estructurada (carente de definiciones, operaciones y soluciones). La muestra de sujetos fue categorizada en dos grandes grupos: jóvenes (infantil y primaria) y mayores (secundaria, universidad y adultos).

Los resultados generales indicaron que la cooperación es mas efectiva en comparación con la competición para la resolución de tareas ( $d=0,55$ ,  $DE=0,79$ ,  $n=63$ ) y a la vez, los miembros que participaron en equipos cooperativos superaron el desempeño competitivo en los cuatro tipos de tareas a resolver: lingüística, no lingüística, estructurada y no estructurada ( $d=0,37$ ,  $0,72$ ,  $0,52$  y  $0,60$ , respectivamente). Estas diferencias se manifestaron igualmente a favor del aprendizaje cooperativo en todas las edades y en los estudios de baja, media y alta calidad metodológica.

Específicamente, en lo que respecta al grupo de mayores –secundaria, universidad y adultos–, todas la tareas obtuvieron una magnitud del efecto media–alta. Sin embargo, no encontraron diferencias significativas entre estudiantes de infantil y primaria cuando se compara con estudiantes de secundaria, universidad y adultos ( $F(1, 59)=0,27$ ,  $p < .61$ ) (ver tabla 11)

Tabla 11: Resultados por tipo de tarea y nivel educativo del meta-análisis de Qin et al. 1995

Tipo de tarea	Infantil y primaria			Secundaria y adultos		
	ME	DS	N	ME	DS	N
Total	0,49	0,74	28	0,60	0,85	33
Lingüística	0,26	0,82	12	0,40	0,80	18
No Lingüística	0,67	0,64	16	0,83	0,87	15
Estructurada	0,34	0,82	18	0,67	0,66	23
No estructurada	0,77	0,49	10	0,43	1,22	10

Fuente: Qin, Johnson &amp; Johnson (1995, p. 135)

Cuando se analiza los dos grupos de edad –jóvenes y adultos– con tareas lingüísticas y no lingüísticas ( $F(1, 57)=0,58, p < .45$ ) y tareas estructuradas y no estructuradas ( $F(1, 57)=0,29, p < .59$ ), las diferencias no fueron significativas. Tampoco se encontraron diferencias en variables como el tipo de publicación ( $F(3, 59)=0,35, p < .79$ ), asignación al azar ( $F(2, 60)=0,98, p < .38$ ) y tiempo de duración del programa ( $F(3, 57)=1,52, p < .22$ ). En cambio, si hallaron diferencias significativas en cuanto a la década de publicación, entre los años 1900 y 1980 ( $F(5, 57)=4,36, p < 0,002$ ) pero el número reducido de estudios antes de los años 1960 resta relevancia a esta conclusión porque el 80% de estos se concentra entre 1970 y 1980.

Concluyen que la superioridad del aprendizaje cooperativo se evidencia con más fuerza en las tareas no lingüísticas, como por ejemplo, la resolución de problemas matemáticos y visuales–espaciales como figuras geométricas, rompecabezas y laberintos en comparación con aquellas que requieren de la solución de problemas lingüísticos en donde los estudiantes discuten las posibles respuestas a preguntas a través del lenguaje oral y escrito. Las razones del por qué el aprendizaje cooperativo es más efectivo en este tipo de tareas, se debe al intercambio de información, la generación de variedad de estrategias para la solución de problemas, el incremento de la habilidad para transferir el enunciado a las ecuaciones y el desarrollo de una representación cognitiva compartida del problema.

En esta misma línea de investigación, estudiar si el grupo facilita el rendimiento, las actitudes y el autoconcepto de estudiantes de primaria, secundaria y universidad así como cuáles son los factores que explicarían la variabilidad de los estudios, qué tipos de agrupamientos son los mejores y bajo qué condiciones, ha sido el objetivo general del estudio de Lou, Abrami, Spence, Poulsen, Chambers & D'Apollonia (1996). El interés por todas estas variables encauzó a estos autores a hacer dos análisis.

El primer análisis comparó el rendimiento con y sin agrupamiento a partir de 103 resultados independientes de 51 estudios con un total de 16.073 alumnos en donde solo siete investigaciones pertenecieron a experiencias en la universidad. La magnitud del efecto global fue baja con una magnitud del efecto de 0,17 ( $IC_{95\%}=+0,16; +0,23$ ) cuando la estrategia utilizada ha sido agrupar a los estudiantes en contraste con situaciones de no agrupamiento.

Son varias las variables moderadoras que explicaron este efecto. El instrumento de medida fue un predictor significativo ( $Q_b(2)=70,02; p <0,5$ ), también los test cuando han sido elaborados por el docente ( $d=0,42; n=16$ ) o por el investigador ( $d=0,34; n=21$ ) en comparación con test estandarizados ( $d=0,07, n=61$ ). El tamaño del grupo ha sido otra variable sustantiva que arrojó datos interesantes porque los grupos pequeños (3–4 miembros,  $n=38$ ) obtuvieron una magnitud del efecto mayor 0,22 en comparación con los grupos grandes 0,11 (8–10 miembros,  $n=24$ ), las parejas 0,15 ( $n=13$ ) y los grupos medianos -0,02 (5–7 miembros,  $n=17$ ).

Estudiantes con bajo, medio y alto nivel de rendimiento se beneficiaron del trabajo en grupo ( $d=0,37, 0,19$  y  $0,28$ , respectivamente) aunque no de manera uniforme porque estudiantes de bajo rendimiento se desempeñaron mas significativamente que los alumnos de nivel medio. Este mismo comportamiento se evidenció con respecto al aprendizaje en todas las asignaturas aunque Matemáticas y Ciencia ( $d=0,20$ ) obtuvieron una puntuación mas alta en comparación con lectura, Arte y otras ( $d=0,13$ ). En cuanto a las etapas educativas, 4to, 5to y 6to grado de primaria ( $d=0,29, n=36$ ) obtuvieron mayores magnitudes del efecto cuando se comparan con secundaria y bachillerato ( $d=0,17, n=25$ ), con 1ero, 2do y 3er grado de primaria ( $d=0,08, n=30$ ) y con la universidad 0,19 ( $IC_{95\%}= -0,03; +0,42, Q_w 8,55, n=7$ ).

El segundo análisis, se centró en la comparación entre grupos heterogéneos y homogéneos a partir de 20 resultados independientes de 12 estudios seleccionados, con la diferencia de que Lou y colaboradores (1996) no hicieron esta vez, un análisis por subgrupos según la etapa educativa.

Los resultados globales señalaron una magnitud pequeña de 0,12 ( $IC_{95\%}= +0,01; +0,24$ ) en donde la efectividad de los grupos homogéneos es mayor a los heterogéneos no siendo esta diferencia uniforme en todos los resultados ( $Q_t(19)=43,90, p<0,5$ ). El rango de los tamaños individuales osciló entre -1,75 a 1,12 con 13 resultados a favor de los

grupos homogéneos, 6 a favor de los grupos heterogéneos y en 1 no se encontraron diferencias.

Estudiantes con bajo rendimiento lograron un mejor resultado en grupos heterogéneos que en grupos homogéneos ( $d = -0,60$ ), aquellos con un nivel medio se beneficiaron mas de los grupos homogéneos que los heterogéneos ( $d = 0,51$ ) y para los de alto rendimiento no se evidenciaron diferencias entre ambos grupos ( $d = 0,09$ ). En Matemáticas y Ciencias, el efecto de la composición del grupo no fue diferente de cero ( $d = -0,00$ ) mientras que en lectura, los grupos homogéneos son mas eficientes en contraste con los grupos heterogéneos ( $d = 0,36$ ).

Aprender en grupo facilita el aprendizaje y así lo demostró la investigación de Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996). Sus resultados ponen de manifiesto que el trabajo en grupo dentro de clase facilita el aprendizaje del estudiante a diferencia del no agrupamiento, especialmente en Matemáticas y Ciencia. Además los grupos pequeños conformados por 3 o 4 miembros fueron más efectivos. Los estudiantes con un nivel bajo se beneficiaron más en grupos heterogéneos, al contrario que los estudiantes con un nivel de conocimiento promedio, quienes obtuvieron mayor ventaja si el grupo era homogéneo. Concluyen estos autores que la mejor agrupación de clase es aquella que promueve, mediante la distribución de los estudiantes, métodos y materiales para el aprendizaje en pequeños grupos.

Este mismo grupo, a partir del estudio realizado en 1996, con Lou como investigador principal, en el año 2000, en conjunto con sus colegas Abrami y Spence, analizaron 103 resultados independientes provenientes de 51 estudios desde primaria hasta la universidad a través de dos modelos ponderados de regresión múltiple.

El objetivo principal del estudio fue desarrollar un modelo de parsimonia de los factores que inciden en la variabilidad de los resultados de los efectos en el rendimiento de los estudiantes según el tipo de agrupamiento. Codificaron variables agrupadas en tres grandes categorías: medición de la variable dependiente, metodológicas y sustantivas. Dos investigadores independientes calcularon la magnitud del efecto y la calidad de los estudios siendo el acuerdo interobservador de 72,90% y 88,24%, respectivamente.

La magnitud global del efecto fue positivo pero bajo con  $+0,16$  ( $IC_{95\%} = +0,13; +0,19$ ) en el rendimiento a favor del agrupamiento sobre la instrucción a toda la clase, valor que no difiere del estudio previo reportado por estos mismos autores ( $d = 0,17$ ).

En el primer modelo de regresión simple, en donde se analizaron las 17 variables del estudio previo de Lou et al (1996), cinco variables fueron las que explicaron los resultados: la formación de los profesores (5%), el tipo de instrucción de los grupos pequeños (6%), especificidad del agrupamiento (2%), nivel educativo (5%) y la habilidad de los estudiantes (14%).

En el segundo modelo, de las 24 variables codificadas, el 48% del total de la varianza se explicó a través de los instrumentos de medida (12%), la formación de los profesores (16%), especificidad del agrupamiento (22%), tipo de instrucción para el trabajo en pequeños grupos (28%), el nivel educativo (34%) y el nivel de habilidad de los estudiantes (48%).

Los hallazgos en este modelo indicaron que los efectos del agrupamiento se evidencian con mas fuerza cuando: (a) los instrumentos de medida son elaborados por los propios investigadores en comparación con los test estandarizados, (b) la formación que recibieron los profesores quienes trabajaron con grupos pequeños es diferente a quienes trabajaron con toda la clase, (c) el método aplicado es el aprendizaje cooperativo como la mejor forma de agrupamiento, (d) el rendimiento es la condición para estructurar los grupos y no cuando se considera el sexo o la cohesión grupal, (e) el aprendizaje en pequeños grupos se lleva a cabo para estudiantes con altos niveles de funcionamiento en comparación con aquellos que cuentan con un nivel bajo de habilidades y (f) el efecto del agrupamiento es mas fuerte en estudiantes de primaria en comparación con estudiantes universitarios cuando los otros factores han sido controlados.

Ante esta realidad, los autores Lou, Abrami y Spence (2000) se plantearon varias hipótesis que inducen a la reflexión del quehacer en el contexto universitario. La primera, es que en primaria tanto profesores como estudiantes utilizan con mayor frecuencia estrategias de aprendizaje grupal cooperativo en comparación con la universidad. Los grupos cooperativos en Educación Superior tienden a ser menos estructurados porque existe un aprendizaje previo en la dinámica de esta metodología o debido a la madurez de los estudiantes. La segunda, obedece a que los estudiantes universitarios están mucho más acostumbrados a estudiar en clases magistrales y en consecuencia tienen una actitud menos positiva hacia el trabajo cooperativo. La tercera y última, una división desigual de las responsabilidades ocurre con mas frecuencia cuando no se cuenta con una estructura organizada de la tarea a realizar. Esta situación contribuye a que los estudiantes universitarios prefieran las clases magistrales y no el

trabajo en pequeños grupos. Concluyen los autores afirmando que no todos los agrupamientos son igualmente efectivos.

Sin embargo, el estudio realizado por Lou et al. (1996) no se vio exento de críticas. Desde Europa, Prais (1999) comentó que en las escuelas americanas al ser lo suficientemente grandes, el método que se utiliza es agrupar a los alumnos en clases paralelas o en función de sus habilidades, de esa manera, el profesor no se encuentra bajo presión para dividir su tiempo entre grupos distintos dentro de cada clase.

En cambio, en las escuelas en el continente europeo, se apunta a una serie de estrategias mas profundas para mantener a toda la clase en conjunto al mismo ritmo. Abrami, Lou, Chambers, Poulsen y Spence (2000) respondieron a las críticas de Prais (1999) argumentando que sus evidencias demostraron la efectividad de esta metodología en el rendimiento y en distintas habilidades de los estudiantes aunque coincidieron con Prais en que la gran variabilidad de los estudios necesita continuar siendo investigada y que la utilización solo de esta metodología no mejora el aprendizaje.

Stanne, Johnson y Johnson (1999) con la finalidad de comparar el aprendizaje cooperativo, competitivo e individual en cuanto a rendimiento, relaciones interpersonales y salud psicológica, revisaron 754 estudios entre los años 1900 y 2009, de los cuales, un 41,4% se han llevado a cabo con estudiantes universitarios y adultos: 278 estudios con edades comprendidas entre 19 y 22 años (36,9%) y 34 investigaciones con estudiantes mayores de 23 años (4,5%).

Encontraron que el rendimiento promedio de una persona en una actividad cooperativa está por encima dos tercios de desviación estándar, en comparación con el rendimiento promedio de una persona en una actividad competitiva ( $d=0,67$ ) e individual ( $d=0,64$ ). Cuando solo se incluyen aquellos estudios con alta validez interna, aumenta la magnitud en la competitividad 0,88 pero se mantiene en situaciones individuales con 0,61.

En las experiencias cooperativas las estrategias de alto nivel cognitivo y razonamiento moral se promueven con mayor frecuencia en comparación con situaciones individuales ( $d=0,97$ ), resultado que es igualmente mayor cuando el tiempo invertido en tareas cooperativas se compara con individuales ( $d=1,17$ ).

Comentan que luego de un periodo de relativo obscurantismo en los años 60, el aprendizaje cooperativo ha florecido y hoy en día es aplicado en escuelas y universidades en distintas partes del mundo y en áreas diferentes desde infantil hasta



programas universitarios y para adultos, lo que ha contribuido a validar la teoría y clarificar las definiciones conceptuales (Stanne, Johnson & Johnson, 1999).

En años más recientes, en el 2000, Bowen decidió medir el efecto del aprendizaje cooperativo cuando se compara con el aprendizaje individual en estudiantes de bachillerato y universitarios quienes estudian Química. Un total de 437 estudiantes de bachillerato y 1.110 de universidad participaron en las 15 investigaciones, de las cuales, 11 solo se llevaron a cabo en la universidad y 4 de ellas provienen del meta-análisis Springer, Stanne y Donovan (1999).

De los 30 resultados independientes, la magnitud del efecto global del rendimiento en Química fue de 0,37 (DE=0,39) a favor del aprendizaje cooperativo. Visto el importante número de publicaciones sobre los efectos del aprendizaje cooperativo en estudiantes desde primaria hasta bachillerato y que en general, la mayoría de estos estudios se han llevado en las áreas de Ciencias y Matemáticas, Bowen (2000) recomendó utilizar esta metodología para la enseñanza de la Química lo que implica un necesario cambio en el currículo y un mayor esfuerzo por parte del profesorado en las situaciones de enseñanza universitarias.

Johnson, Johnson y Stanne (2000) afinaron el rumbo de los meta-análisis al estudiar específicamente la eficacia de los métodos de aprendizaje cooperativo mas comunes al compararlos con situaciones de competitividad e individualismo. Un total de 158 estudios entre los años 1970 y 1990 cumplieron con los criterios de inclusión y de estos, cuatro provenían del sureste asiático, tres del Medio Oriente, tres de Europa, cuatro de África y el resto de Norteamérica. Encontraron estudios desde primaria hasta la universidad, inclusive con población adulta. No obstante, el número de documentos primarios en Educación Superior (n=33) y con adultos (n=4) fue muy bajo en comparación con el número de investigaciones en otras etapas educativas, por ejemplo primaria con 73 documentos.

Específicamente, el aprendizaje cooperativo favoreció el rendimiento cuando se compara con el aprendizaje individual en los métodos *Learning Together*, *Academic Controversy*, *Group Investigation*, *Teams–Games–Tournaments*, *Team Assisted Individualization*, *Student Teams Achievement Divisions*, *Cooperative Integrated Reading and Composition* y *Jigsaw*. Este beneficio también se evidenció en todos los métodos cooperativos en contraste con situaciones competitivas (ver tabla 12).

Tabla 12: Resultados métodos cooperativos del meta-análisis de Johnson Johnson &amp; Stanne (2000)

Métodos cooperativos	Cooperativo/individual				Coop/comp
	ME	DE	N	IC <sub>95%</sub>	ME
Learning Together (LT)	1,03	0,69	56	±0,08	0,85
Academic Controversy (AC)	0,91	0,59	11	±0,19	0,67
Group Investigation (GI)	0,62		1	±0,86	0,37
Teams–Games–Tournaments (TGT)	0,58	0,43	5	±0,22	0,48
Team Assisted Individualization (TAI)	0,33	0,26	8	±0,12	0,25
Student Teams Achievement Divisions (STAD)	0,29	0,71	14	±0,14	0,51
Cooperative Integrated Reading and Composition (CIRC)	0,18	0,00	1	±0,22	0,18
Jigsaw	0,13	0,29	5	±0,21	0,29

Fuente: adaptación de Johnson, Johnson &amp; Stanne (2000)

Aunque es evidente que todos los métodos cooperativos mejoraron significativamente el rendimiento de los estudiantes cuando se confronta con el aprendizaje individual, no es de olvidar, el reducido número de estudios en universidades que cumplieron con los criterios de inclusión y que además, este meta-análisis tampoco realizó un análisis por subgrupo según la etapa educativa.

Son los propios investigadores del estudio quienes advierten que “las diferencias de la magnitud del efecto entre los distintos métodos tienen que interpretarse con cautela” (Johnson, Johnson & Stanne, 2000, s.p.) debido a que los instrumentos de medida no fueron equivalentes. Esto supone como muy bien señalan los autores, futuras líneas de investigación en donde el criterio de calidad metodológica se asuma como otra variable moderadora a analizar.

Por último, con el desarrollo de la informática y su inclusión en ambientes educativos Lou, Abrami y D’Apollonia (2001) asumieron el reto de estudiar los efectos sociales, el desempeño en las tareas y el rendimiento cuando los estudiantes aprenden a usar la tecnología, ya sea a través de pequeños grupos o de manera individual en documentos publicados entre los años 1966 y 1999. El meta-análisis incluyó 486 resultados independientes de 122 estudios primarios, en donde participaron en total 11.317 estudiantes, de los cuales solo 65 investigaciones midieron rendimiento en el nivel universitario.

Los resultados indicaron que el trabajo en pequeños grupos fue significativamente positivo en comparación con el aprendizaje individual, tanto en el rendimiento individual de cada miembro del grupo ( $d=0,16$ ,  $n=178$ ,  $IC_{95\%}=+0,12$ ;  $+0,20$ ), como en el rendimiento grupal ( $d=0,31$ ,  $n=39$ ,  $IC_{95\%}=+0,20$ ;  $+0,43$ ).

Aunque el estudio no calculó específicamente la magnitud del efecto en función de la etapa educativa, esta osciló en los estudios primarios en Educación Superior entre -0,04 y +3,37 y, en cuanto a las asignaturas: Arte -0,04 y +0,99 (n=10); Estudios Sociales +0,26, habilidades en Tecnología -0,04 y +1,40 (n=25); Educación -0,01 y +0,43 (n=7); Lengua extranjera 0,08, Medicina 0,12, Geografía -0,37 y +0,53 (n=5); Biología +0,48 y +0,61 (n=2); Estadística +0,12 y +0,33 (n=2); Salud +0,45 y 3,37 (n=6); Ingeniería +0,17 y 0,08 (n=2); Administración 0,00 (n=4); Química +0,41; Matemáticas +0,81; Ciencias -1,14 y +0,03 (n=2).

Los factores contextuales y pedagógicos moderadores en el rendimiento individual durante el aprendizaje en pequeños grupos fueron: la experiencia o pautas en el grupo de trabajo, las estrategias de aprendizaje grupal, la asignatura, el tipo de programa, el tipo de publicación y el nivel de habilidades. Así, el efecto del aprendizaje en grupos pequeños es significativo cuando: (a) los estudiantes tienen experiencias en el trabajo en grupo o en seguir instrucciones, (b) se utilizan estrategias específicas de aprendizaje cooperativo, (c) el tamaño del grupo es pequeño (por ejemplo dos miembros), (d) usan tutoriales o software para prácticas o lenguajes de programación, (e) las asignaturas son relativas a Informática y Ciencias Sociales y por último, (f) el nivel de los estudiantes se situaba en niveles extremos (bajos o altos) en función de la capacidad requerida.

Las estrategias que utilizaron los docentes para promover el aprendizaje cooperativo también fueron un factor pedagógico importante, porque influyeron en cómo los estudiantes aprenden a trabajar en pequeños grupos usando la tecnología. Los autores aseguraron que no todas las experiencias son positivas cuando uno de los miembros del grupo realiza solo el mínimo esfuerzo. El estudio confirmó cómo el aprendizaje cooperativo se ha perfeccionado con la inclusión de la tecnología al favorecer las habilidades sociales y el rendimiento académico de los estudiantes (Lou, Abrami & D'Apollonia, 2001).

### **2.2.3. Meta-análisis centrados solo en la universidad**

El aprendizaje cooperativo es un método viable, con un gran recorrido de investigación pero todavía subutilizado en el nivel universitario (Jones & Jones, 2008; Paulson & Faust, 2008) en donde prevalece la clase magistral “The majority of college faculty still teach their classes in the traditional lecture mode” (Faust y Paulson, 1998, p. 3) lo que favorece el aprendizaje individual en detrimento de otras metodologías activas de

enseñanza. Tan es así, que son exiguos los meta-análisis dedicados únicamente al estudio del aprendizaje cooperativo en el ámbito universitario.

El meta-análisis de Springer, Stanne y Donovan (1999) es una prueba de ello. El objetivo de esta revisión cuantitativa se centró en el estudio del rendimiento, la retentiva y las actitudes. Los criterios de inclusión para la selección de estudios primarios fueron: (a) estudiantes universitarios del programa SMET, es decir, que cursaban las asignaturas de Ciencias, Matemáticas, Ingeniería o Tecnología (esta última no entendida como usuarios de tecnología sino aquellos cursos relacionados con la Salud) en instituciones norteamericanas, (b) grupos cooperativos dentro o fuera de clase conformados entre 2 y 10 estudiantes, (c) estudios que se han llevado a cabo en aulas ordinarias y no en laboratorios o aulas especialmente diseñadas para la investigación, (d) estudios publicados a partir del año 1980 y (e) con suficiente información estadística para el cálculo de la magnitud del efecto.

La revisión de la literatura reportó 383 documentos de los cuales solo 39 cumplieron con los criterios de inclusión y de ellos, la mayoría, 37 (94,9%) presentó información sobre rendimiento, solo 9 (23,1%) sobre retentiva y 11 (28,2%) vinculados a las actitudes. Demostraron la eficacia del aprendizaje cooperativo en el rendimiento académico ( $d=0,51$ ) cuando se compara con situaciones no cooperativas (ver tabla 13)

Tabla 13: Resultados rendimiento del meta-análisis de Springer, Stanne & Donovan (1999)

Medidas		Estudios	Resultados		Magnitud del efecto		$Q_i$
		N	N	Estudiantes	No ponderada	Ponderada	
Rendimiento	Muestras independientes	37	49	3472	0,51	0,51	90,10*
	Muestras no independientes		116		0,44	0,44	250,50*

Fuente: adaptado de Springer, Stanne & Donovan (1999, p. 38)

Todas las magnitudes ponderadas del efecto han sido estadísticamente significativas (el 95% del intervalo de confianza no incluye el cero). En magnitudes del efecto no ponderadas no se calculó el nivel de significancia

\* $p < 0,05$

Entre los factores que explicaron la variabilidad de la magnitud del efecto en el rendimiento se encontraron variables metodológicas como el papel del investigador. En aquellos estudios en donde el investigador fue el responsable del estudio, la magnitud del efecto se incrementó ( $d=0,73$ ) en comparación con aquellos en donde el investigador no participó directamente en la instrucción de la actividad ( $d=0,41$ ). Estudios con diseño de dos muestras con grupo control y experimental reportaron un

mayor efecto ( $d=0,57$ ) en contraste con aquellos con pre-test y post-test en una sola muestra ( $d=0,30$ ) (ver tabla 14).

Tabla 14: Resultados variables moderadoras del meta-análisis de Springer, Stanne & Donovan (1999)

Medidas		Estudios	Muestras independientes		Magnitud del efecto		$Qb$	$Qw$
		N	N	Estudiantes	No ponderada	Ponderada		
Investigador	Investigador principal	15	18	1261	0,73	0,73	14,32*	30,32*
	Otro	12	18	1305	0,37	0,41		15,04
Diseño de investigación	Una muestra	6	12	764	0,42	0,30	9,03*	19,11
	Dos muestras	31	37	2559	0,54	0,57		61,95*
Tipo de institución	Cuatro años	30	41	3163	0,57	0,54	6,70*	76,22*
	Dos años	6	7	276	0,15	0,21**		7,15
Disciplina	Ciencias	9	14	1071	0,46	0,42	3,85	23,59*
	Matemáticas	22	29	1956	0,52	0,53		46,25*
	Ciencias de la Salud	6	6	445	0,55	0,66		16,41*
Tipo de publicación	Revistas	21	29	2166	0,57	0,56	2,94	46,81*
	Otros	16	20	1306	0,42	0,43		40,34*

Fuente: adaptado de Springer, Stanne & Donovan (1999, p. 39)

A excepción de la puntuación con dos asteriscos todas las magnitudes ponderadas del efecto han sido estadísticamente significativas (el 95% del intervalo de confianza no incluye el cero). En magnitudes del efecto no ponderadas no se calculó el nivel de significancia

\*  $p < 0,05$

Sin embargo, no todas las variables metodológicas se asociaron con las diferencias medias del efecto. Por ejemplo, no se encontraron diferencias en cuanto a las disciplinas. Los cursos vinculados al área de Salud (Fisioterapia y Enfermería) obtuvieron una magnitud del efecto de 0,66, comparado con 0,53 en Matemáticas (Estadística y Tecnología) y 0,42 en Ciencias (Química, Biología y Física). Tampoco se encontraron diferencias significativas en función del tipo de publicación. Los artículos de revistas científicas lograron una ligera puntuación mayor ( $d=0,56$ ) en contraste con conferencias en congresos ( $d=0,43$ ) (ver tabla 14).

En cuanto a la conformación de los grupos, con solo 48 muestras independientes, no se encontraron diferencias significativas en el rendimiento en aquellos grupos predominantemente de mujeres ( $d=0,39$ ) en comparación con grupos mixtos ( $d=0,55$ ). Tampoco encontraron diferencias significativas en estudiantes que escogieron

itinerarios con y sin menciones ( $d=0,61$ , para ambos), maestros en formación ( $d=0,40$ ) y alumnos del primer curso ( $d=0,52$ ) y de otros cursos ( $d=0,54$ ) (ver tabla 15)

Tabla 15: Resultados variables moderadoras del meta-análisis de Springer, Stanne & Donovan (1999)

Medidas		Estudios	Muestras independientes		Magnitud del efecto		$Qb$	$Qw$
		N	N	Estudiantes	No ponderada	Ponderada		
Conformación	Predominantemente mujeres	8	13	737	0,41	0,39	3,50	26,42*
	Heterogéneos	28	35	2653	0,54	0,55		57,44
Titulación	SMET con mención	10	11	1243	0,65	0,61	4,35	33,23*
	SMET sin mención	5	8	435	0,62	0,61		4,63
	Maestros en formación	6	11	601	0,48	0,40		20,60
Curso	1er curso	12	15	1417	0,52	0,52	0,01	31,79
	Otros cursos	7	10	766	0,58	0,54		32,40

Todas las magnitudes ponderadas del efecto han sido estadísticamente significativas (el 95% del intervalo de confianza no incluye el cero). En magnitudes del efecto no ponderadas no se calculó el nivel de significancia.

\*  $p < 0,05$

Fuente: adaptado de Springer, Stanne & Donovan (1999, p. 41)

Otras variables relacionadas con el aprendizaje en grupos pequeños, como sesiones complementarias impartidas fuera de clase obtuvieron una mayor magnitud del efecto en el rendimiento ( $d=0,65$ ) en comparación con aquellas sesiones impartidas en clase ( $d=0,44$ ). Aunque no se encontraron diferencias significativas en la asignación de sujetos y tampoco entre el tiempo invertido en la tarea y el rendimiento, sí se evidenció una tendencia a un mejor desempeño cuando el tiempo para ejecutar la tarea es mediano ( $d=0,73$ ), es decir, un semestre –entre una semana y once semanas– en comparación con un tiempo muy alto ( $d=0,52$ ) –a partir de once semanas– o bajo ( $d=0,52$ ) –cuando se utilizan los grupos cooperativos para tareas puntuales–. Por último, la magnitud del efecto en el rendimiento en trabajo en grupo fue mayor cuando la medida son exámenes o pruebas ( $d=0,59$ ) que cuando se trata de instrumentos estandarizados ( $d=0,33$ ) (ver tabla 16)

Tabla 16: Resultados variables moderadoras del meta-análisis de Springer, Stanne & Donovan (1999)

Medidas		Estudios	Muestras independientes		Magnitud del efecto		$Qb$	$Qw$
		N	N	Estudiantes	No ponderada	Ponderada		
Sesiones	En clase	26	34	2223	0,48	0,44	6,86*	51,57*
	Clases complementarias	9	13	1090	0,60	0,65		30,09*
Asignación	Azar	9	13	573	0,46	0,46	2,04	11,07
	No-azar	7	7	451	0,67	0,65		13,39*
	Auto-selección	4	5	306	0,50	0,59		4,59
Tiempo	Alto	12	13	1168	0,53	0,52	3,98	24,05*
	Medio	8	10	515	0,63	0,73		7,88
	Bajo	7	10	538	0,52	0,52		12,83
Evaluación	Examen	31	40	2614	0,56	0,59	10,90*	65,51*
	Test estandarizados	8	13	1011	0,37	0,33		39,15*

Todas las magnitudes ponderadas del efecto han sido estadísticamente significativas (el 95% del intervalo de confianza no incluye el cero). En magnitudes del efecto no ponderadas no se calculó el nivel de significancia.

\*  $p < 0,05$

Fuente: adaptado de Springer, Stanne & Donovan (1999, p. 42)

Springer, Stanne y Donovan (1999) concluyeron que el efecto del aprendizaje cooperativo en el rendimiento es robusto. Son los propios autores quienes reconocen que la magnitud del efecto encontrado en este meta-análisis excedió positivamente los resultados de estudios anteriores invitando a seguir trabajando en el aprendizaje en pequeños grupos con estudiantes universitarios.

Es esta conclusión tan alentadora la que ha sorprendido a quienes investigan el proceso de enseñanza-aprendizaje en la formación de futuros médicos, porque en la mayoría de las investigaciones en las Ciencias Médicas el efecto del aprendizaje cooperativo ha sido pequeño o negativo cuando se ha empleado como metodología, lo que ha guiado a Colliver, Feltovich y Verhulst en el año 2003, a hacer una revisión crítica del meta-análisis de Springer, Stanne y Donovan del año 1999 con la finalidad de encontrar las mejores evidencias para el currículo de Medicina. Hallaron deficiencias sustantivas y metodológicas importantes que atañen a la calidad de los estudios primarios.

De los 37 estudios originales del meta-análisis de Springer y colaboradores seleccionaron solo nueve por haber escogido a los sujetos al azar en donde se cumplen las condiciones de grupo control y experimental –Dees, 1991; Hanshaw, 1982; Iwasiw

& Goldenberg, 1993; Kacer, Rocklin & Weinholtz, 1992; Lovelace & McKnight, 1980; Lynch, 1984; Reglin, 1990; Smith, 1984; Urion & Davidson, 1992–.

De estos estudios, cinco no cumplen con la definición conceptual de aprendizaje cooperativo desde los planteamientos de Slavin y Johnson y Johnson –Hanshaw (1982), Kacer, Rocklin & Weinholtz (1992), Lovelace & McKnight (1980), Lynch (1984) y Reglin (1990)– y los otros cuatro restantes comparan el aprendizaje en pequeños grupos con el individual –Dees (1991), Iwasiw & Goldenberg (1993), Smith (1984), Urion & Davidson (1992)– donde en tres no hay efecto, uno es negativo y otro positivo.

Para Colliver, Feltovich y Verhulst (2003) “the meta-analysis considered here does not support the application of small-group learning in medical education and it raises questions about meta-analysis in education with implications for evidence-based education” (pp. 4–5).

La otra y última evidencia más reciente –hasta la fecha– que demuestra el escaso número de meta-análisis dedicado al estudio de situaciones individuales, competitivas y cooperativas en el nivel universitario es la revisión cuantitativa de Johnson y Johnson, el año 2002a, en donde el aprendizaje cooperativo es un ejemplo de la teoría validada por la investigación aplicada a la práctica docente.

Un total de 312 estudios fueron comparados para medir la eficacia de situaciones cooperativas, competitivas e individuales en la universidad bajo tres grandes categorías: rendimiento, calidad de las relaciones y salud psicológica.

Las investigaciones seleccionadas que conformaron el estudio datan entre los años 1910 y 2000, en donde el mayor número se ubica en la década de los 80 (30,1%) seguida por los años 70 (20,2%), 60 (19,6%) y 90 (17,9%). Esta metodología se aplicó en asignaturas muy diferentes como Ciencias, Ciencias Sociales, Tecnología, Inglés, Lectura, Matemáticas, Psicología, Salud, Educación Física y en una variedad de tareas que requieren habilidades verbales, Matemáticas y Física. Mientras que la mayoría de los estudios fueron de Norteamérica otros en menor proporción procedieron de Europa, Medio Este y Asia. Prevalcieron los artículos de investigación (81,1%), en donde se aplicó la metodología en un tiempo de duración entre 1 y 9 sesiones (79,8%) y la asignación de sujetos al azar (48,1%) aunque también un número alto de estudios se categorizaron como no aleatorizados (38,8%) (ver tabla 17).



Tabla 17: Características de los estudios del meta-análisis de Johnson & Johnson (2002a)

Características	Número de estudios	Porcentaje
<b>DÉCADA</b>		
1910–19	1	0,3
1920–29	5	1,6
1930–39	5	1,6
1940–49	2	0,6
1950–59	17	5,4
1960–69	61	19,6
1970–79	63	20,2
1980–89	94	30,1
1990–99	56	17,9
2000–09	8	2,6
<b>ASIGNACIÓN</b>		
Asignación azar sujetos	150	48,1
Asignación azar grupo	41	13,2
No asignación	121	38,8
<b>TIPO DE PUBLICACIÓN</b>		
Artículo de investigación	253	81,1
Libro	2	0,6
Tesis de Maestría	3	1,0
Tesis Doctorales	27	8,7
Reportes técnicos	17	5,4
No publicados	10	3,2
<b>DURACIÓN (SESIONES)</b>		
1–9	249	79,8
10–19	15	4,8
20–29	13	4,2
30–39	10	3,2
40–49	14	4,5
50–99	11	3,4
<b>Total</b>	<b>312</b>	<b>100</b>

Fuente: Johnson & Johnson (2002a, p. 122)

De los 312 estudios solo 168 midieron rendimiento y de estos no se ofrece información de la magnitud del efecto con respecto a posibles variables moderadoras. Los resultados indicaron que el desempeño se favorece cuando se aplica el aprendizaje cooperativo en comparación con el aprendizaje competitivo ( $d=0,49$ ) y con situaciones individuales ( $d=0,53$ ) en donde las tareas que mas se beneficiaron con el uso de esta metodología son las verbales (lectura, escritura y lenguaje oral), aquellas relacionadas con las Matemáticas y con habilidades físicas (natación, golf y tenis).

Los autores resaltaron la necesidad de estudios rigurosos en la universidad que mejoren la práctica educativa a través del conocimiento sistemático de métodos de enseñanza

University teaching has not changed significantly for hundreds and perhaps thousands of years [afirmación que habría que matizar]. The many attempts to reform university teaching have often demonstrated positive effects but have then

been discontinued. One explanation for the resistance of teaching to change is that instructors fail to apply the same scientific rigor to their teaching as they do to their research. Professors as scientists and intellectuals typically ask for proof when a colleague presents a scientific conclusion, yet when it comes to what constitutes good teaching, professors often accept uncontested folklore and mythology. Many of the recommendations made about teaching, furthermore, are based more on stories and promising ideas rather than conclusions from rigorous research. What is lacking is the successful application of theory and research to instructional methods (Johnson & Johnson, 2002a, p. 119).

Por tanto, el aprendizaje cooperativo es un método legitimado porque ayuda a los estudiantes a aprender (Cohen, 1994a), tal y como lo demostró la síntesis realizada recientemente por Hattie (2009) sobre más de 800 meta-análisis relacionados con el rendimiento en distintos niveles educativos.

El estudio de esta investigadora relativo a seis grandes dominios –estudiante, casa, escuela, currículo, profesor y enfoques de enseñanza–, mostró que los resultados no solo favorecen el aprendizaje cooperativo cuando se compara con el individual ( $d=0,59$ ), sino que ubica a esta metodología de enseñanza en el puesto 24 de 138 variables moderadoras estudiadas, con una magnitud del efecto positiva de 0,59 versus el aprendizaje competitivo en el puesto 37 ( $d=0,54$ ) y muy lejos del aprendizaje competitivo versus el individual en el puesto 97 ( $d=0,24$ ). Encontró que el aprendizaje cooperativo es superior a la competencia en todos los grupos de edad pero, que los resultados mas fuertes se evidencian en primaria y secundaria en contraste con la universidad.

Si bien es cierto, la síntesis realizada por Hattie (2009) no se focalizó solo en el nivel universitario, es verdad que los resultados demostraron cómo entre el grupo de estrategias centradas en la enseñanza, el aprendizaje cooperativo cuando se compara con el aprendizaje individual ( $d=0,59$ ) se ordena en el ranking de variables moderadoras por debajo de *Reciprocal Teaching* ( $d=0,74$ ) y de *Problem-Solving Teaching* ( $d=0,61$ ) en el mismo puesto que otras estrategias de enseñanza como *Direct Instruction* ( $d=0,59$ ) e inclusive muy alejada de *Inductive Teaching* ( $d=0,33$ ), *Inquiry Based Teaching* ( $d=0,31$ ) y *Problem-Based Learning* ( $d=0,15$ ).

Resulta también llamativo y a la vez confirmatorio que entre las primeras variables que mejoran el rendimiento, se encuentran aquellas asociadas al profesor, como es el *Micro Teaching* ( $d=0,88$ ), *Teacher Clarity* ( $d=0,75$ ), *Teacher-Student Relationships* ( $d=0,72$ ), *Not Labeling Students* ( $d=0,61$ ) y *Profesional Development* ( $d=0,61$ ) porque es el

profesor y su formación una pieza angular –pero no la única– en la mejora de la productividad y el desempeño de los estudiantes.

Slavin (1980, 1989a, 1992a), en numerosos escritos, reconoce la larga tradición del estudio del rendimiento como resultado de la eficacia del aprendizaje cooperativo, al mismo tiempo que admite que son contadas las investigaciones experimentales en el nivel universitario y que lamentablemente muchas de ellas presentan variedad de problemas metodológicos pero que, al margen de estas limitaciones, hay evidencias suficientes que avalan los efectos positivos en el rendimiento aunque los resultados no sean tan consistentes como en primaria y secundaria (Fraser et al. 1977 y Sherman & Thomas, 1986 en Slavin, 1989a).

Es indiscutible que “La necesidad de utilizar el aprendizaje cooperativo en el aula universitaria viene motivada, no sólo por las funciones de la universidad, sino también como alternativa metodológica a modelos más directivos e individualistas” (León del Barco y Latas, 2007, p. 274) en donde “The field of higher education can benefit considerably from the more frequent use of quantitative meta-analysis as a tool for understanding the existing literature” (Bowman, 2012, p. 380).

En definitiva, si en Educación Superior son dos las únicas síntesis integrativas orientadas solo al estudio del aprendizaje cooperativo frente al aprendizaje individual en el ámbito universitario, la de Springer, Stanne y Donovan en el año 1999 y la última en el año 2002 por los hermanos Johnson, la realización del presente meta-análisis cobra sentido y se justifica (a) por ser el nivel educativo en donde existe un menor número de síntesis cuantitativa cuando se compara con otras etapas educativas, (b) porque ya ha transcurrido a la fecha de hoy 10 años desde la última revisión cuantitativa, (c) porque el estudio de otras variables explicativas podrían dar luces a nuevos hallazgos sobre las condiciones mas óptimas de esta metodología en el rendimiento académico de estudiantes universitarios y por último, (d) porque la demanda de metodologías activas eficaces sustentada en la Evidencia Basada en la Investigación Educativa es un reto permanente para la mejora de la calidad educativa.

## **PARTE EMPÍRICA**



### **Capítulo III. JUSTIFICACIÓN Y OBJETIVOS DEL ESTUDIO**

El aprendizaje cooperativo no es una metodología reciente. La perspectiva histórica lo sitúa, como primer paso de la cooperación intra-aula, en el movimiento humanista francés del siglo XVIII a través de los postulados de enseñanza mutua de Pestalozzi y las propuestas educativas de Rousseau (Ovejero, 1990a) que, con los inicios de la Psicología científica de finales del XIX, desembocó en el funcionalismo de Dewey. De ahí que haya pasado por las aportaciones de la Psicología Social de mediados del siglo XX sin olvidarnos de la influencia específica que han ejercido las teorías de Piaget, Ausbel y Vygotsky, entre otros (Serrano, Pons y Ruiz, 2007).

El inicio histórico del estudio del aprendizaje cooperativo, en la década de los 20 del siglo pasado, es un indicador del largo trayecto que ha recorrido esta metodología de aprendizaje y a su vez, del número importante de investigaciones que se han publicado hasta nuestros días. Comenzó en pequeña escala en laboratorios americanos y se llevó a las aulas de clase en el año 1970. Es en ese momento cuatro grupos de investigación, uno de ellos en Israel y otros tres en Estados Unidos empezaron a aplicar los métodos cooperativos en contextos educativos (Slavin, 1991).

Precisamente, los años 70, es la década que reporta el mayor número de investigaciones centradas en la universidad cuando colegas cercanos a Slavin como los hermanos Roger Johnson y David Johnson, se unieron a esta iniciativa para formar a profesores universitarios (Johnson & Johnson, 1999d).

La experiencia se llevó a cabo en las clases de ciencia. Para ese momento, la formación en aprendizaje cooperativo se focalizó en cinco grandes áreas: (a) resumen y ampliación de la teoría relativa a cooperación y competición, (b) revisión de la investigación existente en orden a validar o rechazar la teoría y establecer lo conocido y desconocido, (c) elaboración de programas de investigación longitudinales para identificar las condiciones bajo las cuales la cooperación, la competitividad y el trabajo individual son efectivos y (d) operacionalizar la teoría en una serie de procedimientos para el uso docente y personal administrativo en universidades y programas de formación (Johnson & Johnson, 1999d).

Los resultados de la experiencia permitieron comprender los límites entre lo que es y no es aprendizaje cooperativo, los diferentes métodos cooperativos, los elementos básicos del trabajo en grupo y en especial, cómo alcanzar resultados positivos cuando la cooperación ha sido estructurada cuidadosamente en las aulas universitarias (Johnson & Johnson, 1999d).

En los años 80, la investigación americana dio un paso más allá. Buscó comprender cuáles eran las variables y los factores implicados en situaciones cooperativas. Básicamente intentaba indagar las causas y mecanismos que explicaban por qué el aprendizaje cooperativo resultaba ser una metodología que ofrecía resultados positivos tanto en el rendimiento académico como en las habilidades sociales (León del Barco, Gozalo y Vicente, 2004).

Atendiendo a esta necesidad, un grupo de investigación de la Universidad de Johns Hopkins, comenzó a diseñar modelos comprensivos de aprendizaje cooperativo, intentando integrar los métodos utilizados por los docentes en el diseño curricular, con la finalidad de comprender la eficacia de esta metodología en áreas académicas centrales o estructurales (Slavin, 1999c).

En esa década, se quiso conocer si los beneficios encontrados con estudiantes universitarios podrían aplicarse también en educación primaria y bachillerato, reduciéndose así el número de experiencias en Educación Superior. De esta forma, los docentes de primaria no solo lograron mejorar los métodos, sino que simultáneamente, construyeron otras innovaciones alrededor de los principios generales de esta metodología. Diseñaron el programa *Success for All*, que fue realmente como su nombre lo indica, un éxito. Cientos de estudios realizados por distintos grupos de investigadores demostraron el triunfo de este programa al que se sumaron otros programas cooperativos para el nivel de Educación Inicial (Slavin, 1999c).

Siguiendo en el tiempo, en los años 90, la investigación se tornó mucho más específica, es decir, se centró en el proceso ¿qué ocurre durante la interacción de los participantes? Y además ¿qué factores previos condicionan la eficacia del aprendizaje cooperativo?

Los estudios analizaron el habla entre los miembros del grupo en cuanto a la naturaleza, cantidad y variabilidad en función de las tareas y tipos de equipos cooperativos (Bennet, 1985; Bennet & Dunne, 1991). Otros, se interesaron en estudiar los procesos cognitivos implicados en la interacción, es decir, el solicitar y dar ayuda (Nelson Le Gall, 1992;

Webb, 1992) o el explicarse a sí mismo y el explicar a otros (Johnson & Johnson, Stanne & Garibaldi, 1990; O'Donnell, Dansereau, Hall, Skaggs, Hythecker, Peel & Rewey, 1990). Con respecto a los factores previos, las investigaciones se focalizaron en el estudio de las diferencias individuales como variable predictora de resultados positivos. Así, se halló cómo la habilidad verbal y el estilo cognitivo (Rewey, Dansereau, Dees, Skaggs & Pitre, 1992) y las habilidades cognitivas de inducción, extraversión y orientación social (Hall, Rocklin, Dansereau, Skaggs, O'Donnell, Lambiotte & Young, 1988) son variables que favorecen un aprendizaje cooperativo eficiente (ver todos estos estudios en León del Barco, Gozalo y Vicente, 2004).

Pero ¿qué pasó con la trayectoria de investigaciones en el ámbito universitario? Johnson, Johnson & Smith (2007) son los primeros en admitir que es a finales de los años 90, cuando el aprendizaje cooperativo regresa nuevamente a la universidad americana después de un cúmulo importante de investigaciones centradas en otros niveles educativos pero, con el inconveniente de ser una estrategia pedagógica bien investigada aunque infrautilizada en las aulas universitarias (Jones & Jones, 2008; Paulson & Faust, 2008).

En España, habiéndose comenzado a estudiar el aprendizaje cooperativo con mayor sistematicidad en los años 90 –mucho mas tarde que en Norteamérica– las experiencias en las aulas universitarias españolas han ido en aumento como resultado de las exigencias del Espacio Europeo de Educación Superior.

Son esas exigencias en materia de innovación e investigación educativa las que tienen como finalidad formar, por una parte, a un estudiante capaz de responder a los retos que le exige la sociedad como ciudadano activo y futuro profesional y, por otra, a un profesor universitario que mejore los procesos de aprendizaje a través de la búsqueda de nuevas estrategias y metodologías docentes, como resultado de una sólida reflexión didáctica (Apodaca, 2009; Rosales, 2009).

El Plan Bolonia, con las fortalezas y debilidades propias de toda propuesta, demanda del profesor universitario una reflexión permanente de sus competencias y de su capacidad para adaptarse a estos nuevos tiempos, condiciones propias que toda innovación exige. Promover la cooperación entre alumnos, entre alumnos y profesores y por qué no, entre profesores y profesores como colegas de trabajo, es uno de los tantos



cambios metodológicos y competencias que se le exige al profesorado en el proceso de enseñanza y aprendizaje.

El trabajo en equipo y el diseño de metodologías y organización de actividades son competencias profesionales de todo profesor universitario (Zabalza, 2003) en donde el dominio del contenido es fundamental siempre que demuestre su “...capacidad estratégica para planificar, implicar y facilitar la consecución de las metas” (p. 331). De allí que

saber de contenidos disciplinares de un campo de conocimiento no basta, por insuficiente, para la demostración de la calidad en el proceso. Ahora lo que se nos pide es diáfano: enseñar y/o facilitar el aprendizaje de las competencias profesionales y académicas. La función docente tendrá que abarcar también el diseño de la actividad de aprendizaje y el del mejor escenario para ejercitar un aprendizaje interactivo y cooperativo (Santos Rego, 2005, p. 8).

Sin pretender ofrecer una única respuesta a una realidad tan compleja, como es el desafío de la innovación educativa, es cierto que el aprendizaje cooperativo, como metodología de aprendizaje cumple con dos condiciones fundamentales: incrementa el rendimiento académico y además ofrece beneficios en el desarrollo de actitudes personales y habilidades grupales.

El trabajar en equipo, el trabajar cooperativamente, es una competencia transversal no solo presente en muchos de los planes de estudios universitarios, sino además, una competencia profesional muy demandada.

Por ejemplo, en América el reporte de *Hart Research Associates* del 2008 así lo demuestra, o las principales publicaciones de la *Organisation for Economic Co-operation and Development* (2005) –en inglés OECD– y Europa junto con otros continentes y países como Latinoamérica, Rusia y África a través del TUNING Educational Structures (2000) valoran el trabajo en grupo como competencia interpersonal.

En España, el Estatuto del Estudiante Universitario exige la incorporación de competencias transversales en las titulaciones (Real Decreto 1791/2010, de 30 de diciembre) o el Real Decreto 1509/2005, de 16 de diciembre, por el que se modifica el Real Decreto 55/2005, de 21 de enero, en donde las competencias están orientadas a formar en la disciplina del título universitario, pero también, en otras competencias de carácter mas amplio asociadas con el empleo. Son los empleadores quienes también

valoran el trabajo en grupo como una competencia necesaria en el mundo laboral (ACCENTURE, 2007).

Los estudiantes universitarios requieren prepararse para su vida profesional en donde la productividad y en particular, el saber trabajar cooperativamente, sumado al conocimiento propio de cada una de las profesiones que desempeñarán, dibujen un perfil de salida que satisfaga las exigencias sociales y del mercado laboral.

Por tanto, el devenir histórico en Norteamérica, aunado a las exigencias del Espacio Europeo de Educación Superior y a las experiencias de innovación educativa en la universidad española, representan un panorama internacional y nacional volcados en la construcción de modelos educativos con metodologías activas de enseñanza, siendo el aprendizaje cooperativo un método legitimado que mejora el rendimiento académico de los estudiantes (Cohen, 1994a).

Prueba de ello es que la literatura científica cuenta con una cantidad representativa y significativa de estudios empíricos en aprendizaje cooperativo. Estos han permitido revisiones sistemáticas y meta-análisis que confirman su efectividad en comparación con otros métodos tradicionales en cuanto a la productividad y actitudes hacia el aprendizaje (Johnson & Johnson, 1989; Slavin, 1995).

Pero, un gran número de esas síntesis cuantitativas, se han llevado a cabo en primaria, secundaria y/o bachillerato y en menor número en la universidad porque “Pure experimental research studies of cooperative learning of some have not been very commonly done at the college/university level” (Baker, 1995, p. 44) siendo el rendimiento, una de las variables que mas se ha analizado pero a la vez, una de las mas difíciles debido a su heretogeneidad operacional (Kyndt et al. 2013).

En etapas pre-universitarias, se ha demostrado la eficacia del aprendizaje cooperativo en el rendimiento, actitudes y relaciones interétnicas (Sharan, 1980); en el rendimiento en Matemáticas (Slavin, Lake & Grooff, 2009); lectura (Slavin, Cheung, Groff & Lake, 2008); y con estudiantes con altas capacidades (Neber, Finsterwald y Urban, 2001).

También se han estudiado los factores internos que inciden en la dinámica de los grupos cooperativos (Abrami, Lou, Chambers, Poulsen & Spence, 2000) y el efecto de las relaciones sociales, el autoconcepto y el método *Peer Assisted Learning* (Ginsburg-Block, Rohrbeck & Fantuzzo, 2006) o de la tecnología cuando se trabaja en grupos pequeños para mejorar la comprensión lectora (Chambers et al. 2011), entre otros.

A pesar de que esta metodología se ha ido consolidando en el tiempo y sus beneficios académicos y sociales han sido comprobados, son escasos los meta-análisis orientados a estudiar los efectos del aprendizaje cooperativo en comparación con el aprendizaje individual en el rendimiento académico de estudiantes universitarios ¿Por qué? Son varias las razones que apuntan en esta dirección sin pretender por ello que sean las únicas.

La primera, el número de estudios centrados en aprendizaje cooperativo es reducido (Springer, Stanne & Donovan, 1999; Baker, 1995) debido a que en la universidad prevalece la clase magistral como metodología de enseñanza lo que favorece a su vez el aprendizaje individual (León del Barco y Latas, 2007; Ministerio de Educación y Ciencia, 2006).

La segunda y quizás causa-consecuencia de la anterior, las dificultades propias de la puesta en práctica de una metodología que requiere de un estudiante acostumbrado a trabajar con otros y de un profesorado con formación en materia de innovación educativa porque “Learning in groups has some disadvantages as well and comes at a ‘cost’. When certain conditions are met, the benefits outweigh the costs and cooperative learning is very effective” (Van der Vleuten y Driessen, 2014).

La tercera y no necesariamente la última, la falta de apoyo institucional en recursos físicos y económicos (León del Barco y Latas, 2007; Santos Rego, 1990) que se une con escasas experiencias cooperativas en “ámbitos educativos que preceden a la universidad (Bachillerato, Formación Profesional)” (Ministerio de Educación y Ciencia, 2006, p. 48).

En consecuencia, las revisiones cuantitativas que miden en distintas etapas educativas el rendimiento en situaciones cooperativas incluyen un pequeño número de investigaciones en la universidad (Bowen, 2000; Johnson & Johnson, 1989; Johnson, Maruyama, Johnson, Nelson & Skon, 1981; Johnson, Johnson & Stanne, 2000; Lou, Abrami & D’Apollonia, 2001; Lou, Abrami, Spence, Poulsen, Chambers & D’Apollonia, 1996; Lou, Abrami & Spence, 2000; Stanne, Johnson & Johnson, 1999). A esta proporción reducida de meta-análisis se añade que son solo dos los que se han realizado exclusivamente con estudiantes universitarios: el primero, en el año 1999 a cargo de Springer, Stanne y Donovan y el segundo, encabezado por los hermanos Johnson en el año 2002a.

La demanda en renovaciones metodológicas y el creciente número de publicaciones científicas en aprendizaje cooperativo –por ser el enfoque estudiado con mayor sistematicidad a diferencia del colaborativo y del competitivo–, reafirman la necesidad de integrar ese conocimiento a través de revisiones críticas y rigurosas que permitan valorar la evidencia científica y es, el meta-análisis, el tipo de revisión de la producción científica capaz de acumular de forma rigurosa y mediante la aplicación de análisis estadísticos los resultados de un conjunto de estudios empíricos sobre un mismo problema de investigación (Glass, 1977; Kavale & Glass, 1981).

Por consiguiente, el **objetivo general** del presente estudio es **comprobar estadísticamente si el aprendizaje cooperativo es eficaz como metodología de aprendizaje cuando se compara con el aprendizaje individual en el rendimiento académico de estudiantes universitarios y de ser eficiente, bajo qué condiciones** siendo los objetivos específicos:

1. Integrar cuantitativamente los resultados del aprendizaje cooperativo y del aprendizaje individual de estudios experimentales y cuasi-experimentales que midan el rendimiento académico de estudiantes universitarios,
2. Analizar las variables moderadoras del aprendizaje cooperativo que favorecen el rendimiento académico de los estudiantes universitarios,
3. Ofrecer orientaciones metodológicas y elementos eficaces para la puesta en práctica del aprendizaje cooperativo en Educación Superior.

La formación del profesorado universitario en técnicas cooperativas es una propuesta interesante para incorporar y por ende, para transformar las actuales metodologías educativas como alternativa a modelos de aprendizaje mas individualistas. El desafío de todo profesional de la docencia es incrementar el rendimiento académico de sus alumnos

As educational researchers, we find ourselves in the mildly embarrassing position of knowing less than we have proven. The proofs reside in a vast literature that is often superciliously scorned and insufficiently respected. Extracting knowledge from accumulated studies is a complex and important methodological problem to which I commend your attention (Glass, 1976, p. 8).



## Capítulo IV. MÉTODO

Las fases propias de todo meta-análisis se resumen en la definición del problema, la revisión de la literatura, los criterios de inclusión/exclusión para la selección de los estudios primarios, el análisis o integración de la evidencia de los estudios primarios, y por último, la interpretación y presentación de la evidencia acumulada (Cooper & Hedges, 1994; Glass, McGaw & Smith, 1981).

Partiendo del problema a investigar que es comprobar estadísticamente si el aprendizaje cooperativo es eficaz como metodología de aprendizaje cuando se compara con el aprendizaje individual y de ser eficiente, bajo qué condiciones mejora el rendimiento académico de estudiantes universitarios, a continuación se presenta y se justifica cada una de estas fases porque para producir un meta-análisis de calidad se requiere “... a deep substantive understanding of the focus area of the review” (Pigott, 2012, p. 143) y a la vez de un plan a priori porque

... can also help avoid the problem of type I errors when reviewers conduct a series of statistical test (...) can also conduct power analyses of the most substantively important test to see how many studies would be needed to detect a given effect (...). An a priori plan identifying potential moderators will also help reviewers handle missing data when it occurs (Pigott, 2012, pp. 144–145).

### ***4.1. Revisión de la literatura***

La validez del meta-análisis dependerá de la exhaustividad con que se han identificado y ubicado los trabajos originales. Siguiendo las recomendaciones de Durlak y Lipsey (1991) se han utilizado: (a) estrategias formales, como la revisión de fuentes primarias y secundarias y (b) estrategias informales, por medio del contacto con autores y grupos de investigación vinculados al área de interés.

#### *Estrategias formales*

Un total de 11 **fuentes primarias** han sido revisadas: 3 bases de datos (*Web of Knowledge* –desde 2014 *Web of Science*–, *Periodicals Archive Online* y *Proquest Dissertation Publishing*); 2 motores de búsqueda (*Google Scholar* y *Google Búsqueda Avanzada*) y 6 revistas especializadas (*European Journal of Social Psychology*, *College Teaching*, *Higher Education Research Development*, *Journal of Educational*

*Psychology, Research in Higher Education Journal* y *Journal College Science*) (ver anexo 3).

La base de datos *Web of Knowledge* tiene en su haber 100 años de resúmenes, mas de 54 millones de referencias que abarcan 5294 publicaciones en Ciencias Sociales en 55 disciplinas, con 760 millones de referencias citadas y 6.5 millones de registros de 157.000 conferencias (hasta el 2012).

La otra, *Periodicals Archive Online* pertenece a *ProQuest*. Se funda en 1938 siendo innovadora para su época al presentar el microfilm como nuevo formato para la preservación de las publicaciones. Dentro de los servicios que ofrece se encuentra *Proquest Dissertation Publishing* que igualmente es otra base de datos exclusiva para disertaciones y tesis doctorales con mas de 2 millones de investigaciones publicadas alrededor del mundo.

Los motores de búsqueda han permitido la recuperación de documentos primarios. *Google Scholar*, rescata literatura académica como artículos, tesis, libros, resúmenes y repositorios online de universidades y asociaciones profesionales, entre otros. *Google Búsqueda Avanzada* busca información en la web que no necesariamente es académica pero a través del filtro se seleccionaron solo aquellos documentos en formato pdf, criterio que acortó la búsqueda y que a la vez “garantizó” información “mas confiable”.

Las revistas especializadas como *European Journal of Social Psychology* ofrece publicaciones europeas en Psicología Social y *Journal of Educational Psychology* en Psicología Educativa mientras que *Higher Education Research Development, Research in Higher Educational Journal, Journal of College Science Teaching* y *College Teaching* son exclusivas en el ámbito universitario.

La otra estrategia formal de búsqueda ha sido las **fuentes secundarias** entendidas como aquellas que contienen a su vez a las fuentes primarias. Éstas facilitan y maximizan el acceso a la información porque presentan un contenido sintetizado y organizado del tema de interés. Se han consultado, un total de 7 fuentes secundarias, a saber: 3 meta-análisis, 3 estudios primarios y una revisión narrativa.

Los tres *meta-análisis* que se han seleccionado incluyen estudios primarios sobre aprendizaje cooperativo en la universidad aunque algunos también han trabajado con otros niveles educativos. Son investigaciones clásicas referenciadas a su vez en numerosos documentos.

Por ejemplo, los hermanos Johnson tienen a su cargo más de una investigación meta-analítica que datan de distintos años y con niveles educativos diferentes. Específicamente, la del año 1989 es relevante vista su dificultad para conseguir el listado de estudios primarios, los cuales se encuentran únicamente en su libro *Cooperation and competition. Theory and research*, caso contrario al meta-análisis mas reciente de estos autores en el año 2002a que no ofrece esta información. También destaca el estudio de Springer, Stanne y Donovan (1999) por ser uno de los primeros llevados a cabo solo con estudiantes universitarios y el de Bowen (2000) porque incluye estudiantes de bachillerato (ver breve descripción de los meta-análisis utilizados en la tabla 18 y el listado de estudios primarios en los anexos 4, 5 y 6).

Tabla 18: Fuentes secundarias consultadas: meta-análisis

Meta-análisis
Autor: Bowen, C. Año: 2000 Nivel educativo: bachillerato y universidad Objetivo: efecto del aprendizaje cooperativo en el incremento del rendimiento y aprendizaje de la química en estudiantes de bachillerato y universitarios N° de estudios: 15 Número de citas en Google: 145 (hasta el año 2012)
Autor: Springer, L. Stanne, M. y Donovan, S. Año: 1999 Nivel educativo: universidad Objetivo: efectos del aprendizaje cooperativo en estudiantes universitarios de Ciencias, Matemática, Ingeniería y Tecnología en cuanto a rendimiento, persistencia y actitud N° de estudios: 39 Número de citas en Google: 908 (hasta el año 2012)
Autor: Johnson, D. y Johnson, R. Año: 1989 Nivel educativo: desde infantil hasta adultos Objetivo: Efectos del aprendizaje cooperativo, competitivo e individual N° de estudios: 529 Número de citas en Google: 2636 (hasta el año 2012)

Fuente: elaboración propia

Asimismo, se han revisado las referencias bibliográficas procedentes de *otros estudios primarios*. Estos se han elegido por ser mucho mas recientes –años 2001, 2008 y 2012– cuando se comparan con los meta-análisis consultados –años 1989, 1999 y 2000– y porque sus autores no pertenecen a los grupos de investigación relevantes en el área de interés como son Slavin, Cohen, Johnson y Johnson, entre otros (ver tabla 19).



Tabla 19: Fuentes secundarias consultadas: otros estudios primarios

Otros estudios primarios
Autor: Onwuegbuzie, A. & DaRos-Voseles, D. Año: 2001 Nivel educativo: universidad Objetivo: Efectos del aprendizaje cooperativo en estudiantes universitarios en la asignatura de métodos de investigación Número de citas en Google: 25 (hasta el año 2012)
Autor: Tanel, Z. & Erol, M. Año: 2008 Nivel educativo: universidad Objetivo: Efectos del aprendizaje cooperativo vs el aprendizaje individual en estudiantes universitarios para la enseñanza del magnetismo Número de citas en Google: 22 (hasta el año 2012)
Autor: Tran, V. & Lewis, R. Año: 2012 Nivel educativo: universidad Objetivo: Efectos del Jigsaw en el rendimiento y memoria de estudiantes universitarios en matemáticas Nº de citas en Google: 1 (hasta el año 2012)

Fuente: elaboración propia

La única *revisión narrativa* ha sido la de Goodsell, Maher y Tinto (1992). Presenta referencias bibliográficas en la universidad organizadas en cuatro categorías que refieren a: qué es el aprendizaje cooperativo, cómo se implementa, cómo se evalúa y dónde utilizarlo. Al final del documento muestra un listado de investigaciones acompañadas de una breve descripción sobre el programa cooperativo aplicado.

Finalmente, los **descriptores** que se han aplicado en las distintas fuentes primarias y secundarias han sido *Cooperative Learning* y *Collaborative Learning* no combinándose con sinónimos como *university*, *college*, *undergraduate*, *higher education* porque al no contar algunas bases de datos con un tesoro, estos términos no permitían filtrar aquellos estudios propios de este nivel educativo corriendo el riesgo de dejarse fuera investigaciones en Educación Superior. De esta forma, se atendió esta limitación aunque el número de revisión de documentos primarios aumentó significativamente.

Es de aclarar que la única fuente primaria en donde se utilizó el descriptor *Cooperative Learning* y no *Collaborative Learning* ha sido en la base de datos *Web of Knowledge* porque históricamente es el término mas utilizado en esta metodología y el que cuenta con un mayor número de investigaciones (Cuseo, 1992) aunque sea el segundo descriptor el que se asocia con el ámbito universitario (Bruffee, 1995).

### *Estrategias informales*

Un total de 60 correos electrónicos se enviaron a autores y grupos de investigación vinculados al área de interés con “el objeto de paliar los posibles efectos nocivos del

fenómeno del sesgo de publicación, [por lo que] se recomienda complementar el uso de fuentes formales con fuentes informales de búsqueda, tales como contactar con expertos de reconocido prestigio en el campo para solicitarles estudios no publicados” (Sánchez–Meca, 2010, p. 56). Sin embargo, no se obtuvo respuesta siendo una limitación a considerar en esta fase del meta-análisis.

#### **4.2. Criterios de inclusión/exclusión**

Los criterios de inclusión establecidos son el resultado de la revisión de meta-análisis clásicos sobre aprendizaje cooperativo a cargo del grupo de investigación de Slavin y colaboradores (Cheung y Slavin, 2012; Cheung & Slavin, 2013; Johnson & Johnson, 2002b; Johnson, Maruyama, Johnson, Nelson & Skon, 1981; Slavin, 1995; Slavin, Cheung, Groff & Lake, 2008; Slavin, Lake, Davis & Madden, 2011) y también de otros autores relevantes que se han especializado en este método de investigación en otras áreas de conocimiento (Cohen, Kulik & Kulik, 1982; Kulik, Schwalb & Kulik, 1982):

1. Idioma: estudios escritos en inglés o en español.
2. Temporalización: estudios desde enero de 1980 hasta junio del año 2012 (ambos inclusive).
3. Nivel educativo: estudiantes universitarios de cualquier carrera que cursan estudios de licenciaturas, grados, diplomados o postgrado (especializaciones, máster y/o doctorados).
4. Tipos de intervención (VI): estudios en donde las variables independientes han sido el aprendizaje cooperativo y el aprendizaje individual. En el aprendizaje cooperativo los alumnos trabajan juntos en grupos reducidos para alcanzar metas compartidas, maximizando así el aprendizaje individual y el de los demás (Johnson & Johnson, 1999b, 1999c). En cambio, en el aprendizaje individual, los estudiantes no requieren vínculos de interdependencia con los otros compañeros para alcanzar los objetivos. Se logra la meta establecida independientemente del esfuerzo de los demás (Johnson & Johnson, 1999b, 1999c).
5. Medición de la variable de resultado (VD): estudios en donde el rendimiento académico –en inglés *student achievement* y otros sinónimos como *student performances*, *success*, *outcomes*– se define como el desempeño o la productividad de un grupo en distintas áreas de conocimiento medido a través de

diferentes instrumentos cuantitativos. En palabras de Roseth, Johnson y Johnson (2008) se entiende como “... performance on a task, with measures including comprehension, quality and accuracy of answers on tests, quality and accuracy of problem solving, frequency and quantity of desired outcome, higher level reasoning and critical thinking, creativity...” (p. 229).

6. Grupo control y grupo experimental: estudios con al menos un grupo control inactivo (aprendizaje individual) y un grupo experimental (aprendizaje cooperativo).
7. Medidas pretest/posttest: estudios que contemplen medidas antes y después o solo después de la intervención.
8. Cálculo de la magnitud del efecto: estudios que cuenten con la suficiente información estadística para el cálculo de la magnitud del efecto.

Por tanto, un total de 3744 estudios aparecieron registrados bajo los descriptores *Cooperative Learning* y *Collaborative Learning* en las fuentes primarias y secundarias seleccionadas. De estos, se han encontrado 2536, de los cuales 2446 han sido excluidos y solo 90 han cumplido con los criterios de inclusión (ver tabla 20)

Tabla 20: Número total de estudios provenientes de la revisión de la literatura

Revisión de la literatura	Nº de estudios			
	Estudios registrados	Estudios potenciales encontrados	Estudios excluidos	Estudios incluidos
Fuentes primarias	2960	1980	1927	53
Fuentes secundarias	784	556	519	37
Total	3744	2536	2446	90

Fuente: elaboración propia

#### 4.2.1. Fuentes primarias

Las fuentes primarias que han sido consultadas suman en su haber 2960 estudios, de los cuales, 1980 han sido encontrados, 1927 se han excluido y solo 53 han cumplido con los criterios de inclusión (ver tabla 21)

Tabla 21: Número total de estudios excluidos e incluidos de las fuentes primarias

Fuentes primarias*		Nº de estudios												
		Estudios registrados	Estudios excluidos no encontrados	Estudios potenciales encontrados	Estudios excluidos por	Estudios potenciales	Estudios excluidos investigación documental	Estudios potenciales empíricos	Estudios excluidos no cumplen definición aprendizaje cooperativo	Estudios potenciales en aprendizaje cooperativo	Estudios excluidos no universidad	Estudios potenciales de universidad	Estudios excluidos no cumplen otros criterios inclusión	Estudios incluidos en el meta-análisis
Bases de datos	Web of Knowledge	1382	654	728	404 repetidos + 12 no escritos en inglés o español + 48 documentos incompletos	658	342	316	20	296	87	209	184	25
	Periodicals Archive Online	70	12	58		53	25	28	0	28	15	13	12	1
	Proquest Dissertation Publishing	40	14	26		28	0	28	0	28	0	28	20	8
Motores de búsqueda	Google Scholar y Google Búsqueda Avanzada	389	24	365		242	82	160	4	156	59	97	83	14
Revistas especializadas	European Journal of Social Psychology	12	0	12		10	1	9	2	7	1	6	6	0
	College Teaching	229	47	182		147	25	122	36	86	0	86	86	0
	Higher Education Research Development	26	16	10		10	0	10	0	10	0	10	9	1
	Journal of Educational Psychology	726	202	524		314	20	294	192	102	74	28	25	3
	Research in Higher Educational Journal	72	10	62		50	18	32	24	8	1	7	6	1
	Journal of College Science Teaching	14	1	13		4	0	4	0	4	0	4	4	0
Total		2960	980	1980		464	1516	513	1003	278	725	237	488	435

\* en todas las fuentes primarias los descriptores han sido *Cooperative Learning* y *Collaborative Learning* a excepción de la base de datos *Web of Knowledge* en donde solo se ha utilizado *Cooperative Learning*

Fuente: elaboración propia

El diagrama de flujo reflejado en la ilustración 3 resume el proceso de inclusión/exclusión de las fuentes primarias, en donde de los 2960 potenciales estudios, 980 no se encontraron, 12 se han excluido porque no estaban escritos en inglés o español, 404 porque estaban repetidos, 48 por estar incompletos, 513 por ser investigación documental, 278 porque aunque han sido recuperados bajo los descriptores *Cooperative Learning* o *Collaborative Learning* no trabajan con esta

metodología de enseñanza–aprendizaje, 237 porque no se han llevado a cabo en la universidad y 435 no acatan otros criterios de inclusión. Finalmente 53 son los estudios que cumplen con los criterios de inclusión (ver ilustración 3)

Ilustración 3: Diagrama de flujo inclusión/exclusión de estudios de fuentes primarias



Fuente: elaboración propia

#### 4.2.2. Fuentes secundarias

Continuando con el proceso de inclusión y exclusión de posibles estudios potenciales para el cálculo de la magnitud del efecto, se han registrado 784 documentos provenientes de las distintas fuentes secundarias consultadas, de los cuales solo se han encontrado 556 y de estos, 519 se han excluido y apenas 37 han cumplido con los criterios de inclusión (ver tabla 22).

Tabla 22: Número total de estudios provenientes de las fuentes secundarias

Fuentes secundarias*		N° de estudios			
		Estudios registrados	Estudios potenciales encontrados	Estudios excluidos	Estudios incluidos
Meta-análisis	Bowen 2000	583	419	390	29
	Springer, Stanne & Donovan 1999				
	Johnson y Johnson 1989				
Otros estudios	Onwuegbuzie y DaRos-Voseles 2001	165	120	113	7
	Tanel & Erol 2008				
	Tran & Lewis 2012				
Revisiones narrativas	Goodsell, Maher & Tinto 1992	36	17	16	1
<b>Total</b>		<b>784</b>	<b>556</b>	<b>519</b>	<b>37</b>

\* en todas las fuentes secundarias los descriptores han sido *Cooperative Learning* y *Collaborative Learning*

Fuente: elaboración propia

### Meta-análisis

El número de estudios procedentes de la revisión de todos los meta-análisis ha sido 583 de los cuales solo se han encontrado 419 documentos y de estos, 29 han cumplido con los criterios de inclusión (ver tabla 23).

La diferencia tan abismal entre el número de estudios registrados, encontrados y centrados en el nivel universitario obedece a varias razones.

La primera, aunque los meta-análisis de Johnson y Johnson (1989) y Bowen (2000) ofrecieron el listado de los estudios primarios utilizados, se tratan de investigaciones muy antiguas que las propias bases de datos no las tienen en su haber requiriendo así la ayuda de otras bibliotecas nacionales e internacionales para su localización. Por razones de costo, la búsqueda de estos documentos a través del servicio de préstamo bibliotecario limitó continuar la pesquisa de un número mayor de documentos.

La segunda, de estos dos meta-análisis, un número importante de estudios primarios quedaron excluidos porque solo se trabajó con aquellos que se implementaron en la universidad –28 en Johnson y Johnson (1989) y 5 en Bowen (2000)–.

Este comportamiento no se presentó en el meta-análisis de Springer, Stanne y Donovan (1999) porque se aplicó solo con estudiantes universitarios, motivo que a la vez explica el número tan elevado de estudios que sí cumplieron con los criterios de inclusión (ver tabla 23).

Tabla 23: Número total de estudios incluidos/excluidos de la revisión de meta-análisis

Meta-análisis	N° de estudios								
	Estudios registrados	Estudios excluidos no encontrados	Estudios potenciales encontrados	Estudios excluidos repetidos	Estudios potenciales encontrados	Estudios excluidos no universidad	Estudios potenciales en universidad	Estudios excluidos no cumplen otros criterios inclusion	Estudios incluidos
Johnson & Johnson 1989	529	150	379	9	374	346	28	19	9
Bowen 2000	15	4	11		9	4	5	5	0
Springer, Stanne & Donovan 1999	39	10	29		27	0	27	7	20
Total	583	164	419	9	410	350	60	31	29

Fuente: elaboración propia

*Otros estudios*

Los tres estudios primarios exigieron la revisión de un total de 165 referencias bibliográficas, de las cuales, un número elevado fueron excluidas por tratarse de referencias documentales y otras por no acatar los criterios de inclusión. Solo 7 estudios si cumplieron con estos criterios siendo necesario resaltar que 5 de ellos provienen de la reciente investigación de Tran y Lewis (2012) (ver tabla 24).

Tabla 24: Número total de estudios incluidos/excluidos de la revisión de otros estudios primarios

Otros estudios primarios	Estudios registrados	Estudios excluidos no encontrados	Estudios potenciales encontrados	Estudios excluidos repetidos	Estudios potenciales encontrados	Estudios excluidos porque son libros	Estudios potenciales	Estudios excluidos no cumplen otros criterios inclusión	Estudios incluidos
Onwuegbuzie & DaRos–Voseles 2001	52	8	44	2	44	28	16	15	1
Tanel & Erol 2008	39	28	11		10	5	5	4	1
Tran & Lewis 2012	74	9	65		64	27	37	32	5
Total	165	45	120		118	60	58	51	7

Fuente: elaboración propia

*Revisiones narrativas*

De la única revisión narrativa elegida para su revisión, la de Goodsell, Maher & Tinto (1992), conformada por un total de 36 referencias bibliográficas, solo se encontraron 17 y de esas, 6 se descartaron por tratarse de libros a los que no se tuvieron acceso y 9 no

cumplieron con los criterios de inclusión quedando solo 1 estudio para su inclusión (ver tabla 25).

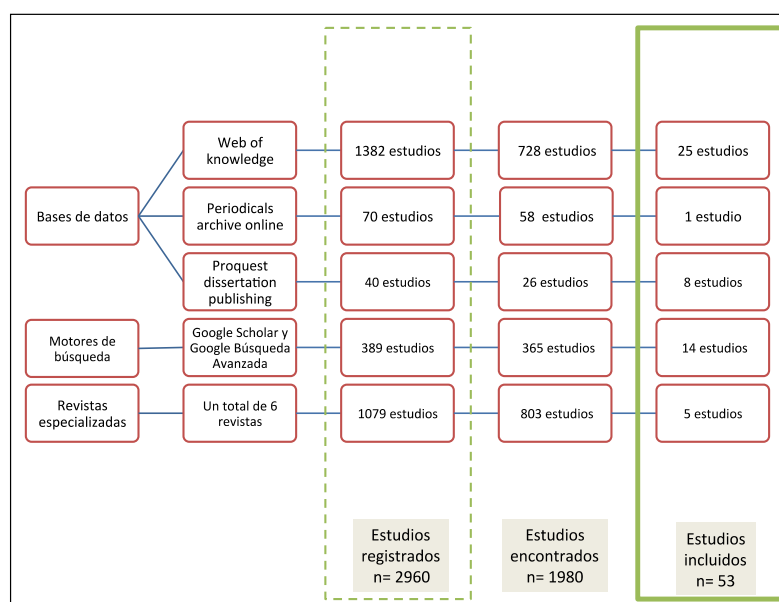
Tabla 25: Número total de estudios excluidos/incluidos de la revisión narrativa

Revisiones narrativas	Nº de estudios								
	Estudios registrados	Estudios excluidos no encontrados	Estudios potenciales encontrados	Estudios excluidos repetidos	Estudios potenciales encontrados	Estudios excluidos porque son libros	Estudios potenciales	Estudios excluidos no cumplen otros criterios inclusion	Estudios incluidos
Goodsell, Maher & Tinto 1992	36	19	17	1	16	6	10	9	1

Fuente: elaboración propia

Finalmente, las ilustraciones 4 y 5 resumen, respectivamente, el número de estudios registrados, encontrados e incluidos de las fuentes primarias y secundarias consultadas y recuerdan que un total de 90 estudios conformaron la muestra final de documentos primarios a ser analizados estadísticamente.

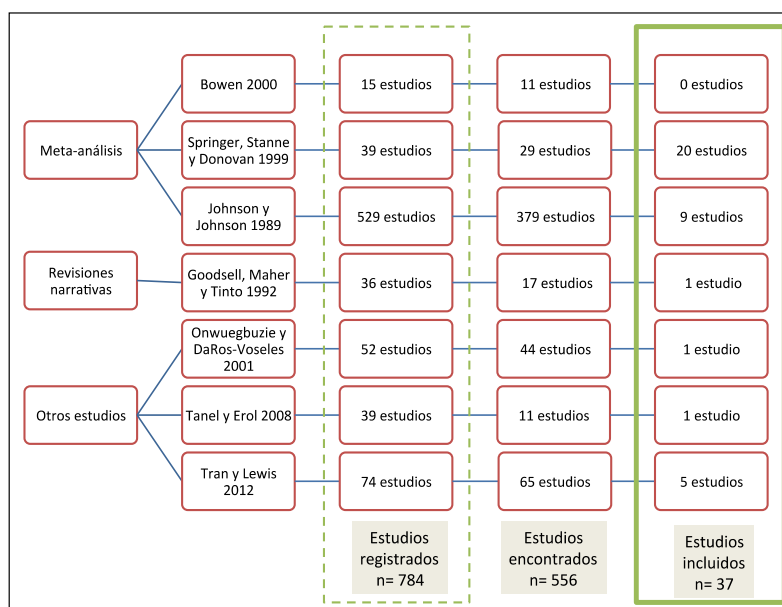
Ilustración 4: Número total de estudios excluidos/incluidos de las fuentes primarias



Fuente: elaboración propia



Ilustración 5: Número total de estudios incluidos/excluidos de las fuentes secundarias



Fuente: elaboración propia

### 4.3. Codificación de los estudios

Este proceso implica la definición y codificación de las características de los estudios siendo tres los criterios que han guiado esas definiciones. El primero “... the characteristic is claimed to have a functional relationship with the phenomena being studied as indicated in the extant literature [el segundo] ...the characteristic is likely to be reported by the experimenter” (Kavale & Glass, 1981, p. 533) y el tercero, la propuesta de Lipsey (1994) para la calificación de las características en extrínsecas, ecológicas y metodológicas.

#### 4.3.1. Característica de los estudios

Como muy bien dicen Durlak & Lipsey (1991) es imposible especificar todas las variables que deberían codificarse en un meta-análisis, lo clave es “Based upon prior research reports, theory and methodological insight, the meta-analyst should be able to enumerate those variables that appear most important for a particular research area and then attempt to code for each” (p. 303).

Por tanto, son 16 las variables contextuales, ecológicas y metodológicas que describen las características de los estudios, las cuales han sido combinadas para su posterior

comprobación por ser consideradas como posibles variables moderadoras en los resultados.

Las *variables extrínsecas* son aquellas que no se relacionan directamente con el objetivo del estudio y se asocian a aspectos vinculados a la identificación y publicación del documento “Se caracterizan por ser externas al propio desarrollo de la investigación, aunque en ocasiones pueden afectar a los resultados de los estudios y poner en evidencia la existencia de deficiencias, sesgos o artefactos en un campo de estudio” (Sánchez-Meca, 2003, p. 102).

1. Años de publicación: ordenados por lustros y décadas.
2. Tipos de documento
  - Publicado: reporte técnico (convenciones, conferencias...), disertaciones (tesis de maestría o doctorales), publicaciones profesionales (libros), publicaciones científicas (artículos revistas científicas)
  - No publicado
3. Lugar de procedencia: continente de las instituciones de los autores que han participado en el estudio (América, Asia, Europa, África u Oceanía).

Las *variables ecológicas* también llamadas variables sustantivas son “... propias del objeto de investigación e incluyen aspectos tales como las características demográficas de las muestras de sujetos sometidas a estudio, cómo se operativizaron las variables, el contexto social, cultural, geográfico, económico, etc. en el que se hizo la investigación, etc.” (Sánchez-Meca, 2003, p. 101). Tienen que ver con el tratamiento, los participantes y el contexto (Sánchez-Meca, 2010).

4. Áreas de conocimiento: agrupación de las asignaturas en donde se ha aplicado el aprendizaje cooperativo según la ciencia de procedencia en función de la Clasificación Internacional Normalizada de la Educación (CINE) propuesta por la UNESCO (2011) para los campos de Ciencia y Tecnología (para ver mas detalles ir al anexo 7)
  - Educación: Formación del Personal Docente y Ciencias de la Educación.
  - Humanidades y Artes: Artes y Humanidades.
  - Ciencias Sociales, Educación Comercial y Derecho: Ciencias Sociales y del Comportamiento, Periodismo e Información, Educación Comercial y Administración y Derecho.

- Ciencias: Ciencias de la vida (a excepción de Medicina y Veterinaria), Ciencias Físicas, Matemáticas y Estadísticas e Informática.
  - Ingeniería, Industria y Construcción: Ingeniería y profesiones afines, Industria y Producción y Arquitectura y Construcción.
  - Agricultura (incluye Veterinaria).
  - Salud y Servicios Sociales: Medicina (Medicina, Servicios Médicos y Enfermería) y Servicios Sociales (Asistencia Social y Trabajo Social).
5. Conformación de los grupos: estructuración de los grupos en homogéneos o heterogéneos en función del nivel de rendimiento académico y/o sexo de los miembros.
6. Número de integrantes: agrupaciones frecuentes de los grupos cooperativos
- Parejas
  - Tres miembros
  - Entre cuatro y seis miembros (ambos inclusive).
7. Duración de la intervención: tiempo desde que comienza hasta que termina la aplicación del programa como
- Sesiones
  - Entre 1 y 4 semanas
  - Entre 5 y 8 semanas
  - Entre 9 y 12 semanas
  - Entre 13 y 16 semanas
  - 17 semanas o mas
8. Clasificación de los métodos cooperativos: las técnicas de aprendizaje cooperativos se han agrupado siguiendo la propuesta de distintos autores (Gavilán y Alario, 2010; Johnson & Johnson, 1994, 1999a, 2002b; Kagan & Kagan, 2009; León del Barco, Gozalo, Felipe, Gómez y Latas, 2005; O'Donnell & Dansereau, 1995; Slavin, 1980, 1985, 1995)
- *Student Team Learning*: STAD, TGT, TAI, CIRC y *Jigsaw II*
  - *Cooperative Investigation*: GI, CO–OP CO–OP, CI y *Jigsaw I*
  - *Dyadic Methods*: *Scripted Cooperation*, *Reciprocal Peer Tutoring*, *Peer Teaching* y *Think–Pair–Share*.
  - *Learning Together*
  - *Cooperative Learning and Computer*: métodos cooperativos que se sustentan en la tecnología para el trabajo presencial cara a cara o a distancia

- *Informal Methods: Spontaneous Group Discussion, Numbered Heads Together, Team Product y Cooperative Review.*
  - Métodos combinados: son aquellos estudios en donde se ha aplicado más de un método cooperativo
9. Estructura de recompensa: en función del método cooperativo, si la recompensa del grupo proviene del rendimiento individual, rendimiento grupal o no hay recompensa
- Recompensas de grupo basadas en los aprendizaje individuales: STAD, TGT, TAI, CIRC y *Jigsaw II*
  - Recompensas de grupo basadas en el producto grupal: LT
  - Sin recompensas de grupo: GI y *Jigsaw I*
10. Estructura de la tarea: en función del método cooperativo, si la estructura de la tarea requiere especialización o no
- Con especialización: *Jigsaw I*, *Jigsaw II* y GI
  - Sin especialización: STAD, TGT, TAI, CIRC y LT
11. Igualdad de oportunidades para la puntuación grupal: en función del método cooperativo, en qué medida el esfuerzo individual contribuye a la puntuación grupal
- Con igualdad de oportunidades: STAD, TGT, *Jigsaw II*, TAI y CIRC
  - Sin igualdad de oportunidades: LT, GI y *Jigsaw I*

Y las *variables metodológicas* que refieren "...al método y diseño de la investigación y, por tanto, se repiten en todos los meta-análisis. Son características metodológicas tales como el tamaño de la muestra, la mortalidad experimental, el tipo de diseño o la calidad del diseño de investigación" (Sánchez-Meca, 2003, p. 101).

12. Asignación: aleatorización de sujetos, aleatorización de grupos, aleatorización de sujetos y grupos, emparejamiento (*matching technique*) y no aleatorización.
13. Situación basal: antes de la intervención, control o no de la situación basal (equivalencia) entre el grupo control y experimental en rendimiento.
14. Efectos del investigador (sesgos): si es el mismo el investigador quien lleva a cabo o no la intervención en el grupo control y experimental
15. Instrumentos de medida: los instrumentos han sido creados por el mismo investigador, son pruebas estandarizadas o validadas o es la combinación de ambos.

16. Calidad metodológica de los estudios: valoración de la calidad metodológica de los estudios primarios (baja, media y alta).

#### **4.3.2. Manual de codificación**

El proceso de codificación de las características de los estudios es un intento por “convert some complex reality to a set of mutually exclusive and exhaustive numbers and will thus always encounter problems of fitting the phenomena to the numerical scheme” (Bullock & Svyantek, 1985, p. 113) porque la falta de fiabilidad en la codificación es una amenaza que atenta contra la validez de las conclusiones del meta-análisis y “esto suele ocurrir especialmente con características complejas que requieren hacer juicios o tomar decisiones por parte del codificador” (Sánchez-Meca y Ato, 1989, p. 629).

La elaboración de un manual de codificación ha permitido atender este problema de validez sin pretender por ello eliminar las dificultades propias del proceso al incluir lo mas exhaustivamente posible las normas y reglas para la codificación de las características de los estudios (ver anexo 8) porque “The meta-analyst should be able to identify all those items of desired information that describe the overall study and collect them in one place in the coding protocol” (Lipsey & Wilson, 2001, p. 76).

Este proceso, se ha llevado a cabo entre dos observadores –uno de ellos el propio investigador del trabajo y el otro una persona experta en meta-análisis– quienes han codificado de manera independiente los primeros 15 estudios primarios, para luego reunirse y revisar el sistema de codificación atendiendo así a los contados desacuerdos que han sido resueltos para la posterior categorización del resto de estudios. De esta manera se han utilizado dos de las estrategias para reducir el error en el proceso de codificación: el entrenamiento de los codificadores y una prueba piloto del manual de codificación (Orwin & Vevea, 1994).

#### **4.4. Calidad metodológica de los estudios: la escala**

Para Durlak y Lipsey (1991) “It is important to keep in mind that there are no absolute standards available to judge design quality or methodological rigor. Different reviewers can and do employ different criteria (...) It is also important to recognize that methodological quality is a multi-dimensional concept” (p. 299).

Asimismo, ambos autores comentan que Glass y sus colegas McGaw y Smith (1981) y otros como Smith (1980) abogan por “criterios liberales de calidad” que ayuden a descubrir empíricamente qué relación existe entre los métodos y las conclusiones de los estudios. Sin embargo, para el grupo de investigación de Slavin (1986) y Wortman y Bryant (1985), los criterios de calidad metodológica se utilizan no como un variable moderadora mas sino todo lo contrario, es un criterio de exclusión de los estudios primarios.

Por tanto, en el presente estudio se ha creado y validado una escala de la calidad metodológica para la valoración de estudios primarios siguiendo la posición de Glass y colaboradores, con la finalidad de comprobar su influencia en los resultados junto con las otras variables metodológicas estudiadas y no como criterio para la inclusión o no de los estudios. La escala se ha elaborado en tres grandes fases: (a) construcción y diseño en español y en inglés, (b) valoración del juicio de expertos y (c) fiabilidad inter-codificadores.

#### **4.4.1. Primera fase: construcción y diseño**

La primera fase de la construcción del instrumento estuvo marcada por la revisión de la literatura, la creación de la versión inicial en español y su traducción al inglés.

Para la construcción inicial del instrumento se revisaron varios documentos especializados en escalas de calidad metodológica en Educación, Psicología y Medicina (Alderson & Green, 2009; Centro Cochrane Iberoamericano, 2012; Chacón, Sanduvete y Alarcón, 2005; Chacón, Sanduvete, Sánchez & Sánchez-Meca, 2007; Coalition for Evidence-Based Policy, 2010; Colorado Department of Education, 2011; CONSORT, 2010; Des Jarlais, Lyles, Crepaz & the Trend Group, 2004; Durlak & Lipsey, 1991; Education Group on Guidelines on Evaluation, 1999; Higgins & Green, 2011; Institute of Education Sciences, 2010; Johns Hopkins University (2014); Lipsey & Wilson, 2001; Ramos-Álvarez y Catena, 2004; Ramos-Álvarez, Moreno-Fernández, Valdés-Conroy & Catena, 2008; Promising practices network, 2012; Sánchez-Meca y Botella, 2010; Shadish, Chacón-Moscoso & Sánchez-Meca, 2005; Shadish & Myers, 2004; Slavin, 1986; Slavin, 2008; STARD, 2003; University of Glasgow, s.f.a; University of Glasgow, s.f.b; U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance, 2003; y por último Valentine, 1994).

Luego de la revisión de la literatura se construyó la escala para medir la calidad de los estudios primarios de diseños experimentales o cuasiexperimentales en función de siete grandes categorías: (a) intervención, (b) conformación de grupos, (c) asignación, (d) muestra, (e) recolección de datos, (f) reporte de resultados y (g) sesgos del investigador. A su vez, cada una de esas categorías estuvo conformada por indicadores para ser evaluados de manera dicotómica. La suma de los distintos indicadores permitiría la obtención de una valoración global de la calidad de los estudios primarios estudiados (ver anexo 9).

Posteriormente, el instrumento fue traducido del español al inglés con la finalidad de que expertos de habla inglesa participaran en su valoración. La estrategia que se ha utilizado para garantizar una adecuada traducción y así minimizar distorsiones en el significado de los ítems, ha sido la selección de un traductor competente en el idioma español e inglés (bilingüe), conocedor de la terminología propia del área de Educación y Psicología y a quien se ha impartido dos horas de formación básica sobre aprendizaje cooperativo, meta-análisis, escalas de calidad metodológica y diseños experimentales y cuasi-experimentales. Durante el proceso, el traductor mantuvo comunicación permanente con el investigador del estudio. Por ejemplo, se consultó si se utilizaba el término “*administer*” o “*apply*” por la palabra “administrar” o “*achievement*” o “*performance*” como traducción de “rendimiento”.

A continuación, la tabla 26 presenta la versión inicial del instrumento en español en donde se observa la estructura y organización de los ítems en siete dimensiones o categorías (para la versión en inglés ir al anexo 10):

Tabla 26: Versión inicial del instrumento de calidad metodológica

I. Intervención: valora la calidad general de la intervención/tratamiento		
	SI	NO
La intervención se sustenta en teoría clara y bien articulada		
Las variables principales del estudio están claramente definidas		
Describe la intervención en cuanto a qué se administra y quién la ha recibido		
Describe cómo la intervención difiere del grupo control		
La intervención se ha llevado a cabo en la universidad o espacios vinculados a la comunidad educativa en donde los propios profesores regulares imparten clases. No se ha llevado a cabo la intervención en aulas especializadas creadas y diseñadas por los investigadores para los propósitos del estudio		
Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural porque el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad		
La intervención puede ser replicada por investigadores quienes tienen distintos niveles de formación		

II. Conformación de grupos: valora la validez interna del diseño utilizado		
	SI	NO
Contempla al menos un grupo experimental y un grupo control		
Hace referencia a si el grupo control no está conformado por sujetos que tenían la opción de participar en la intervención pero se han negado		
El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención		
Se demuestra que no hay diferencias sistemáticas entre el grupo de intervención y de control antes de la intervención (situación basal). Ambos grupos poseen un alto grado de homogeneidad antes de la intervención en cuanto a: niveles de rendimiento académico, datos demográficos...		
El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo		
Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental, es decir, no hay evidencias de contaminación ( <i>cross-over</i> )		

III. Asignación: valora si las secciones/grupos/sujetos han sido asignados aleatoriamente o si, en caso contrario, se ha controlado el influjo de posibles variables que afecten los resultados		
	SI	NO
Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental		
Los grupos han sido asignados aleatoriamente como grupo control o experimental		
Los sujetos han sido asignados aleatoriamente al grupo control o experimental		
No se ha usado la aleatorización para la asignación de sujetos a los grupos pero se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación)		
Advierte sobre cualquier factor que pudiera comprometer el proceso de aleatorización		

IV. Muestra: valora si las secciones/grupos/sujetos han sido asignados aleatoriamente o si, en caso contrario, se han controlado posibles variables que afecten los resultados		
	SI	NO
Describe con detalles los criterios de inclusión de la muestra		
Incluye el número de grupos que han sido escogidos del total de la población para luego especificar el tamaño de la muestra		
El estudio incluye un análisis que asegura que el desgaste de la muestra (si lo hay) no socava la equivalencia de los grupos de intervención y de control. El porcentaje de participantes que se ha perdido durante el proceso de recolección de datos es lo suficientemente pequeño que no afecta los resultados del grupo de intervención y de control (no supera el 80% de la muestra original)		
Se ha solicitado a los sujetos de la muestra su consentimiento para participar en la investigación antes de conocer si estaban asignados al grupo experimental o control		

V. Recolección de datos: valora los instrumentos y medidas aplicadas		
	SI	NO
Los instrumentos utilizados miden con exactitud los resultados reales por los cuales la intervención está destinada a incidir. Existe una correspondencia entre los instrumentos y la intervención realizada		
El grupo control y experimental han sido evaluados con los mismos instrumentos		
Todos los instrumentos aplicados para medir rendimiento han sido validados/estandarizados		
Las medidas de resultado no están a favor del grupo de intervención sobre el grupo control o viceversa. Por ejemplo, un estudio de un programa computarizado para enseñar matemáticas a los jóvenes estudiantes no debe medir los resultados mediante un examen computarizado, ya que el grupo de intervención probablemente tendrá mayor facilidad con el ordenador que el grupo control		

VI. Reporte de resultados: valora los hallazgos encontrados en el estudio		
	SI	NO
Describe cómo la intervención ha afectado los resultados		



Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos		
Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos que solo han obtenido un efecto positivo		
Incluye los resultados principales a los cuales responde la pregunta de investigación así como también resultados secundarios de interés		
Reporta los datos de incluso aquellos miembros del grupo de intervención que no participaron o completaron la intervención ( <i>Intention to treat analysis</i> ) como de aquellos sujetos que terminaron la intervención ( <i>Completers analysis</i> )		
VII. Sesgos del investigador: valora quien administra la intervención		
	SI	NO
El mismo investigador es quien lleva a cabo la intervención en el grupo experimental y control		
Quien aplica la intervención desconoce los objetivos del estudio y en consecuencia a qué grupo pertenece el sujeto que está evaluando (control ciego)		
Valoración global de la calidad del estudio: corresponde a la suma de las puntuaciones en cada una de las siete categorías. Si el indicador ha sido valorado con un NO se puntuará con un valor de 0. Si el indicador ha sido evaluado con un SI se puntuará con un valor de 1. La valoración global de la calidad del estudio oscilará entre 0 pts y 33 pts.		
Categorías	Puntuaciones	
I. Intervención	/7	
II. Conformación de grupos	/6	
III. Asignación	/5	
IV. Muestra	/4	
V. Recolección de datos	/4	
VI. Reporte de resultados	/5	
VII. Sesgos del investigador	/2	
Total calidad global del estudio	/33	

Fuente: elaboración propia

#### 4.4.2. Segunda fase: juicio de expertos

El juicio de expertos permite contrastar en qué medida el instrumento que se ha creado mide adecuadamente las dimensiones de lo que se pretende medir, es decir, si refleja un dominio específico del contenido de lo que se mide. La valoración de los especialistas – y no un índice o coeficiente– es la estrategia que se utiliza para estimar la validez de contenido de un instrumento (Kerlinger, 2002). Son los expertos quienes valoran de manera independiente la “bondad” de los indicadores en cuanto a su relevancia con el universo del contenido, la claridad de la redacción o el sesgo en su formulación (Hernández, Fernández y Baptista, 2010).

En esta fase se ha descrito quienes conformaron finalmente la lista de los expertos; la consistencia interna del instrumento (coeficiente de Alpha de Cronbach) y la valoración cuantitativa (porcentajes) y cualitativa que ha recibido la escala en cuanto a la estructura y organización de las dimensiones, el contenido de los ítems y el sistema de puntuación global.

#### *4.4.2.1. Los expertos*

Los criterios para la selección de los especialistas han sido (a) profesionales o (b) profesores doctores universitarios, (c) en Educación, Psicología, Estadística o Medicina, (d) expertos en al menos una de las siguientes áreas de conocimiento: meta-análisis, aprendizaje cooperativo, métodos de investigación y/o estadística y (e) provenientes de universidades o centros de investigación nacionales e internacionales.

La búsqueda de los expertos se ha realizado de manera intencional. Se ha utilizado la estrategia “bola de nieve” o *snowboling* (Lincoln & Guba, 1985). A partir de una lista inicial de expertos se les solicitó que nombraran a otros que conocieran y que cumplieran con los requisitos de selección. Otra estrategia ha sido la identificación de profesores en universidades españolas privadas y públicas que trabajan en el departamento de métodos de investigación en Educación o en Psicología así como aquellos que pertenecen a grupos de investigación en meta-análisis o que imparten cursos de aprendizaje cooperativo en másters o doctorados.

Específicamente, en la búsqueda de expertos internacionales, se ubicó a grupos de investigación vinculados al aprendizaje cooperativo y a autores que han llevado a cabo estudios experimentales o cuasiexperimentales en aprendizaje cooperativo. También se escribió a especialistas en meta-análisis quienes han publicado libros especializados o bien han impartido conferencias en congresos y jornadas científicas sobre esta área de interés.

A todos ellos –expertos nacionales e internacionales– se invitó a través del correo electrónico a participar de manera voluntaria y anónima en la valoración de la escala de calidad metodológica (ver anexos 11 y 12).

En el ámbito nacional, se contactó inicialmente con

- Un (1) centro de investigación –el Consejo Superior de Investigaciones Científicas (CSIC)– y
- 16 universidades españolas: Centro Universitario Villanueva (adscrito a la Universidad Complutense de Madrid), Universidad Autónoma de Barcelona, Universidad Autónoma de Madrid, Universidad Camilo José Cela, Universidad CEU–San Pablo, Universidad Complutense de Madrid, Universidad de Alcalá, Universidad de Alicante, Universidad de Cantabria, Universidad de Deusto, Universidad de Extremadura, Universidad de Granada, Universidad de Murcia,

Universidad de Sevilla, Universidad de Vitoria, Universidad Internacional de la Rioja y Universidad Nacional de Educación a Distancia.

En el ámbito internacional, se escribió a

- 4 grupos de investigación, tres de ellos especializados en aprendizaje cooperativo *Team Based Learning*, *Success for All Foundation*, *Johns Hopkins University* y uno en meta-análisis *The Campbell Collaboration*.
- 69 universidades de los siguientes países: Australia, Alemania, Bélgica, Canadá, Estados Unidos, Filipinas, Holanda, Hong Kong, Reino Unido, Korea del Sur, Malasia, Nueva Zelanda, Suiza, Taiwan, Turquía y Venezuela.
  - Adventist International Institute of Advance Studies, Anadolu University, Baruch College, Bristol University, Bryn Mawr College, City University of Hong Kong, Columbia University, Concordia University, Durham University, Franklin College Switzerland, George Washington University, Hallym University, Hamdan Bin Mohammed e–University, La Trobe University, Laval University Canadá, Learning Innovations at WestEd, Louisiana State University, Management University, Mohammed e–University, National Taiwan Normal University, North Carolina State University, Northwestern University, Old Dominion University, Peninsula College of Medicine & Dentistry, Purdue University, Rutgers University, Simon Fraser University, U.S. Army Research Institute, Universidad Católica Andrés Bello, Universidad Metropolitana, Universiti Sains Malaysia, University of Alberta, University of Amsterdam, University of Arizona, University of Auckland, University of Bedfordshire, University of Birmingham, University of Buffalo, University of California, University of Cambridge, University of Central Florida, University of Cologne, University of Connecticut, University of Delaware, University of Georgia, University of Groningen, University of Illinois College of Medicine, University of Iowa, University of Kansas, University of Leuven, University of Louisville, University of Loyola, University of Melbourne, University of Michigan, University of Minnesota, University of Montreal, University of New York, University of Northern Iowa,

University of Notre Dame, University of Ottawa, University of Pennsylvania, University of Texas, University of Twente, University of Victoria, University of Wisconsin–Madison, UW–Madison School of Education, Vanderbilt University, Virginia Commonwealth University y Vrije Universiteit.

Finalmente, el número total de expertos ha sido de 35. De estos, solo 30 cumplieron la valoración cuantitativa del juicio de expertos porque los otros cinco solo decidieron hacer comentarios cualitativos que igualmente se sumaron al resto de sugerencias generales:

- 19 especialistas procedentes del ámbito nacional de
  - Un centro de investigación, el Consejo Superior de Investigaciones Científicas (1).
  - 11 universidades españolas como la Universidad de Murcia (1), Universidad de Educación a Distancia (1), Centro Universitario Villanueva (adscrito a la UCM) (2), Universidad de Sevilla (3), Universidad de Granada (1), Universidad Complutense de Madrid (3), Universidad Internacional de la Rioja (1), Universidad de Murcia (1), Universidad CEU–San Pablo (3), Universidad Autónoma de Madrid (1) y Universidad Camilo José Cela (1) y
- 16 especialistas procedentes del ámbito internacional de
  - Un centro de investigación *Johns Hopkins University* (2)
  - 10 universidades como *University of Louisville* (1) y *La Trobe University* (1) en Australia; *University of Leuven* (2) en Bélgica; *University of Texas* (1), *University of Connecticut* (1), *University of Illinois College of Medicine* (1) y *University of Louisville* (1) en Estados Unidos; *Anadolu University* (1) en Turquía y por último, Universidad Metropolitana (4) y Universidad Católica Andrés Bello (1) en Venezuela.

En definitiva, son especialistas en aprendizaje cooperativo, meta-análisis, estadística, métodos de investigación –análisis multivariante, medición, calidad de la investigación en estudios de cohortes y no experimentales, demografía y estadísticas aplicadas– Psicología –Psicología de la Educación, Psicología Educativa e Investigación–, Educación –Didáctica y Organización Escolar e Innovación Docente, Didáctica y

formación en competencias, Diseño Curricular, Educación a distancia, Educación personalizada, evaluación de programas, Investigación Educativa– y Medicina – Métodos Estadísticos para meta–análisis en Salud–.

#### 4.4.2.2. Consistencia interna

Una vez recogidas todas las valoraciones de los expertos, se obtuvo la consistencia interna del instrumento a través del coeficiente de Alpha de Cronbach (ver anexo 13) el cual permite conocer la fiabilidad del instrumento, es decir, en qué medida los ítems de la escala miden un mismo constructo y si correlacionan con un nivel alto entre sí.

Aunque no existe unanimidad entre autores con respecto a cuál es el nivel satisfactorio de fiabilidad porque depende de lo que se quiera hacer con la prueba (Hogan, 2004), se considera que si su uso es para la investigación entonces para Nunnally y Bernstein (1994) una confiabilidad adecuada es de 0,8 mientras que Kaplan y Sacuzzo la ubican en un rango entre 0,7 y 0,8 y Rosenthal en 0,5 pero como regla general se espera que el nivel de confiabilidad no sea menor de 0,8.

En la tabla 27 se observa cómo la escala obtuvo un coeficiente de 0,941, valor que es mayor de 0,9 y se considera como “excelente”. En cuanto a los criterios para la valoración de cada ítem como son redacción, orden, contenido y pertinencia, dos de ellos obtuvieron una consistencia “excelente” con 0,965 (pertinencia) y 0,926 (redacción) y los otros dos un valor de “bueno” con 0,864 (orden) y 0,847 (contenido)

Tabla 27: Coeficiente Alpha de Cronbach para cada criterio de valoración del instrumento

Criterios	Redacción	Orden	Contenido	Pertinencia	Total
Alfa Cronbach	0,926	0,864	0,847	0,965	0,941

Fuente: elaboración propia

Un análisis detallado de las cuatro dimensiones indica que aquellas con un coeficiente mayor a 0,8, valor que se interpreta como “bueno”, corresponde a las dimensiones intervención (0,87), asignación (0,801), recolección de datos (0,845) y sesgos del investigador 0,87. En cambio, las dimensiones que obtuvieron un coeficiente de consistencia interna “aceptable” al ser mayor de 0,7 competen a conformación de grupos (0,762), muestra (0,786) y resultados (0,784) (ver tabla 28).

Tabla 28: Coeficiente de Alpha de Cronbach por cada dimensión del instrumento

Dimensiones	Coeficiente Alpha de Cronbach
I. Intervención	0,87
II. Conformación de grupos	0,762
III. Asignación	0,801
IV. Muestra	0,786
V. Recolección de datos	0,845
VI. Resultados	0,784
VI. Sesgos del investigador	0,876

Fuente: elaboración propia

*4.4.2.3. Comentarios y sugerencias generales*

Con la finalidad de obtener información general de la escala, los expertos respondieron cinco preguntas generales relativas al (a) objetivo del estudio, (b) los valores para la evaluación de los ítems, (c) las instrucciones, (d) el formato y (e) el número de dimensiones que conforman el instrumento siendo los comentarios mas reiterativos y significativos (ver anexo 14):

- Reordenar las dimensiones en función de la validez, interna, externa, de constructo y de conclusión estadística.
- Otorgar un peso diferente a las distintas dimensiones que conforman la escala para la puntuación global.
- Incluir una nueva dimensión que refiera a la validez de conclusión estadística.
- Transformar el instrumento a una escala de Likert sin que ello implique que algunos ítems requieran ser valorados dicotómicamente (si/no).
- Mejorar la instrucción general de la escala en donde queden claros los criterios de inclusión de los estudios primarios.
- Eliminar los ítems que refieren a los criterios de inclusión porque no forman parte de la valoración de la calidad metodológica.
- Mejorar la redacción de los ítems en cuanto a la extensión, disminuir en lo posible la escritura de indicadores en negativo y evitar anglicismos.
- Ampliar la información relativa a la validez de constructo.
- Incluir ítems que permitan valorar tanto estudios experimentales como cuasi-experimentales.

Específicamente, con respecto a la recomendación ofrecida “Transformar el instrumento a una escala de Likert sin que ello implique que algunos ítems requieran ser valorados dicotómicamente (si/no)” algunos indicadores mantuvieron la valoración dicotómica inicial (si/no) mientras que otros fueron cambiados para ser valorados en un rango que

va del 0 al 1, en donde el 0 representa ausencia o una menor calidad metodológica y el 1, una calidad óptima. A su vez, cada uno de estos valores ha sido acompañado de una breve descripción que sirvió como guía para la evaluación de cada ítem, diseño que ha contribuido a su fácil aplicación. Las tablas 29 y 30 representan dos ejemplos.

Tabla 29: Valoración dicotómica de un indicador

	<b>0</b>		<b>1</b>	<b>valor</b>
<b>7.1</b> Selección grupos (población)	No ofrece información sobre cómo han sido seleccionados los grupos (población)	No se han seleccionado aleatoriamente los grupos que van a participar en el estudio	Se han seleccionado aleatoriamente los grupos que van a participar en el estudio	

Fuente: elaboración propia

Tabla 30: Valoración politómica de un indicador

	<b>0</b>	<b>0,5</b>	<b>1</b>	<b>valor</b>
<b>8.1</b> Mortalidad (sujetos del estudio)	El porcentaje de participantes que no ha llegado a terminar el estudio supera el 20% de la muestra original	El porcentaje de participantes que no ha llegado a terminar el estudio es entre 0 y 20% de la muestra original (ambos inclusive)	El porcentaje de participantes que ha iniciado el estudio se mantiene al terminar la investigación, no hay pérdida de sujetos	

Fuente: elaboración propia

#### 4.4.2.4. Las dimensiones y los ítems

Luego de responder los comentarios y sugerencias generales, los expertos puntuaron cada uno de los ítems en un rango del 0 al 2, en donde el 0 representa el menor valor (total desacuerdo con el ítem) y el 2 la puntuación máxima (total acuerdo con el ítem) con respecto a la redacción, el orden (ubicación dentro de la escala), el contenido y su pertinencia en la dimensión.

La tabla 31 presenta la organización de las dimensiones antes y después del juicio de expertos. Siguiendo las recomendaciones de los especialistas (ver comentarios generales) y las propuestas de autores como Cook, Campbell y Peracchio (1990), Shadish, Cook y Campbell (2002) y Valentine (1994), entre otros, las dimensiones de la escala se reorganizaron finalmente en cuatro: validez de constructo, validez interna, validez externa y validez de conclusión estadística.

Tabla 31: Dimensiones de la escala antes y después del juicio de expertos

Antes del juicio de expertos		Después del juicio de expertos	
I. Intervención		I. Validez de constructo	1. Consideraciones teóricas, objetivos e hipótesis del estudio
II. Conformación de grupos			2. Definición de constructos
III. Asignación			3. Administración del tratamiento
IV. Muestra			4. Sesgos del investigador y de los sujetos
V. Recolección de datos		II. Validez interna	5. Diseño investigación
VI. Resultados			6. Selección inclusión/exclusión
VII. Sesgos del investigador			7. Asignación
			8. Mortalidad
			9. Instrumentos de medida
		III. Validez externa	10. Interpretación resultados
			11. Generalización resultados
		IV. Validez de conclusión estadística	

Fuente: elaboración propia

A su vez, los ítems también sufrieron modificaciones en función de las observaciones recibidas.

Las tablas de seguidas presentan solo algunos ejemplos de la versión inicial del ítem (antes de ser evaluado por los expertos), la valoración cuantitativa ofrecida por los especialistas con relación a la redacción, orden, contenido y pertinencia y la versión final del ítem (luego de la valoración de los expertos). Asimismo en estas tablas se han incorporado los comentarios cualitativos los cuales han sido de gran utilidad para la toma de decisiones y las referencias teóricas que han servido de sustento para las decisiones metodológicas asumidas.

Para facilitar la lectura que aparecerán en las siguientes páginas se presenta la versión final del indicador que corresponde solo a la valoración máxima, es decir, aquellos ítems que representan el valor 1 en un rango cuyo valor mínimo es el 0. Se sugiere revisar el anexo 16 para una revisión completa de la descripción que acompaña a cada uno de los valores asignados así como la versión final de la escala ítem por ítem según el juicio de expertos.

En la estructura de la versión inicial de la escala, específicamente en la *dimensión intervención*, un ejemplo de reformulación de ítem en varios indicadores se presenta en el ítem 1.3 “Describe la intervención en cuanto a qué se administra y quién la ha recibido” (ver tabla 32).



El porcentaje obtenido en el orden (100%), la pertinencia (89,7%) y el contenido (81,55%) señalan que es un indicador que ha sido valorado de manera positiva, en cada uno de estos criterios, por más del 80% de los expertos pero un poco más de la mitad (65,5%) consideran que es necesario hacer mejoras en su redacción. La principal observación es que mide más de un indicador a la vez.

Por ello se ha reformulado en tres ítems independientes (ver tabla 32). El primero tiene que ver con la persona que lleva a cabo la intervención “Ofrece información sobre quién es la persona que lleva a cabo la intervención”. Los otros dos, con el tiempo de duración del programa y su distribución en el tiempo “Ofrece información detallada sobre la intervención en cuanto a número, frecuencia y duración de cada sesión” y “El tiempo de duración de toda la intervención es mayor de 12 semanas (inclusive)”

Tabla 32: Versión inicial y final ítem 1.3

Ítem 1.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,9%	8 27,6%	19 65,5%			27 100%		5 18,5%	22 81,5%		3 10,3%	26 89,7%
Versión inicial del ítem			Versión final del ítem			Referencias					
Describe la intervención en cuanto a qué se administra y quién la ha recibido			Ofrece información sobre quién es la persona que lleva a cabo la intervención			CONSORT (2010), Cook & Campbell (1979), Des Jarlais, Lyles, Crepaz & the TREND group (2004).					
			Ofrece información sobre la intervención en cuanto a número, frecuencia y duración de cada sesión			CONSORT (2010), Cook y Campbell (1979), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Morrison, Sullivan, Murray & Jolly (1999).					
			El tiempo de duración de toda la intervención es mayor de 17 semanas (inclusive)			Slavin (2008)					
Redacción: – Dos preguntas en una. – Revisar la redacción en cuanto a cómo se aplica la variable... – Entiendo que se refiera a "quien" la administra. – La utilización del verbo administrar suena a dar algún tipo de sustancias. – Es redundante poner "describe la intervención en cuánto", mejor decir algo como "se especifica qué se administra y quién ha recibido la intervención". – Se describe suficientemente la intervención (qué se administra, cuánto tiempo, quién la ha recibido). – What is "it?" – Describe cómo la intervención es diferente en el grupo control que en el grupo experimental. – Yo eliminaría "quién la ha recibido" puesto que esto corresponde a la categoría "muestra".											

Fuente: elaboración propia

Siguiendo con otro ejemplo, en la *dimensión conformación de grupos*, se ha eliminado el ítem 2.1 “Contempla al menos un grupo experimental y un grupo control”. Aunque si bien es cierto el 80% valora su redacción y el 85,7% su contenido con la puntuación

mas alta, es imprescindible mencionar que se trata de un criterio de inclusión y no de un ítem que mida la calidad metodológica de los estudios primarios (ver tabla 33). El hecho de que algunos expertos se hayan percatado de este error (comentarios cualitativos) ha exigido revisar nuevamente los criterios de inclusión con detenimiento.

Tabla 33: Versión inicial y final del ítem 2.1

Ítem 2.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2	4	24		1	27	2	2	24		1	28
6,7%	13,3%	80%		3,6%	96,4%	7,1%	7,1%	85,7%		3,4%	96,6%
Versión inicial del ítem				Versión final del ítem					Referencias		
Contempla al menos un grupo experimental y un grupo control				Eliminado							
Redacción											
– Parece más un ítem de inclusión/exclusión del meta-análisis.											
– Es imprescindible.											
– Yo cambiaría contempla por existe.											
– Antes de valorar los ítems, eliminar "Conformación de grupos" y apuntar: Validez interna y externa del diseño utilizado: valoración de la aleatorización y las técnicas de control empleadas. Luego, directamente sobre el ítem apunta: se contempla al menos un grupo experimental y un grupo control.											
– ¿no forma parte de los criterios de inclusión?											
– What about if anything is not reported or clearly defined? What about if it meets one but not the other.											
Or if there is contamination that is explained and reported with some numeric information.											
– This would mean that a repeated measures study (with the same group repeated before and after/during a treatment) is getting a low score. It also means that single-case experimental designs would get a low score, although also these designs can be internally valid. Maybe it is therefore better to talk a control and treatment condition, rather than about a control and treatment group.											
– Contempla una división de la "n" en 2 grupos (experimental y control)											
Contenido											
– What if pre and post?											

Fuente: elaboración propia

Pasando ahora a la *dimensión asignación*, los ítems 3.1, 3.2 y 3.3 son un ejemplo de indicadores que se relacionan entre si y que han requerido de la mejora en la redacción.

El ítem 3.1 “Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental” ha sido valorado con la máxima puntuación en la redacción por el 75,9% de los expertos, en el orden con un 92,6%, en el contenido con un 85,2% y en la pertinencia con un 86,2%.

No obstante, para mejorar la escritura y así evitar que se confunda este ítem con el siguiente, el 3.2 “Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental”, se ha cambiado la palabra “aulas (secciones)” por “grupos” quedando redactado de esta manera “Se han seleccionado aleatoriamente los grupos que van a participar en el estudio” (ver tabla 34).

Tabla 34: Versión inicial y final del ítem 3.1

Ítem 3.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,9%	5 17,2%	22 75,9%		2 7,4%	25 92,6%	1 3,7%	3 11,1%	23 85,2%		4 13,8%	25 86,2%
Versión inicial del ítem			Versión final del ítem			Referencias					
Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental			Se han seleccionado aleatoriamente los grupos que van a participar en el estudio			CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos-Álvarez y Catena (2004), Sánchez-Meca, López-Pina y Nuñez-Nuñez (2012).					
Redacción											
– No aplicable sino hay grupo control.											
– Quitar secciones.											
– Estos dos ítems se confunden, pareciera ser el mismo.											
– Para la parte III. Asignación: se sigue valorando la calidad del estudio en cuando a su valor. Puede haber diferencias pero dudo que, así redactado el ítem, el evaluador las pueda captar. Se puede dejar éste pero entonces quitaría el anterior.											
– Modificar la definición de asignación "(...)" que afecten a los resultados".											
– Cambiar la palabra assignation por assignment.											
– This and the next seem redundant. Rephrase.											
– Yo eliminaría este ítem											
Contenido											
– Why classrooms? If individuals, rather that groups or classrooms are assigned to the conditions, the study would get a low score on this topic although assignment of individuals is preferable from a statistical point of view.											

Fuente: elaboración propia

Al haber cambiado la redacción del ítem anterior, la escritura del ítem 3.2 “Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental” no sufrió modificaciones. El hecho de que ambos ítems se relacionen ha sido motivo principal para que la redacción no haya sido valorada tan positivamente con un 76,7%, no así el orden (92,9%), el contenido (85,7%) y la pertinencia (86,2%) (ver tabla 35).

Tabla 35: Versión inicial y final del ítem 3.2

Ítem 3.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
4	3	23	1	1	26	2	2	24	2	2	25
13,3%	10%	76,7%	3,6%	3,6%	92,9%	7,1%	7,1%	85,7%	6,9%	6,9%	86,2%
Versión inicial del ítem			Versión final del ítem			Referencias					
Los grupos han sido asignados aleatoriamente como grupo control o experimental			Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental			CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos-Álvarez y Catena (2004), Sánchez-Meca, López-Pina y Nuñez-Nuñez (2012).					
Redacción											
– ¿Diferencia con el ítem anterior?											
– ¿Qué diferencia hay entre aula y grupo?											
– Uniría los 3 ítems poniendo. Las unidades (aulas, grupos o sujetos) han sido asignados aleatoriamente al grupo control o experimental.											
– Estos dos ítems se confunden, pareciera ser lo mismo.											
– "Me parece igual al ítem 2.6 de la segunda parte".											
– "Parece utilizarse el término "grupo" en 2 sentidos (como grupo de sujetos y como grupo experimental o de control). En todo caso, aunque se entiende qué se quiere decir, la formulación no es clara.											
– Same as above.											
– Yo lo redactaría así: "los tratamientos han sido asignados aleatoriamente a los grupos". Yo colocaría este ítem como el segundo.											
Contenido											
– What is the difference with the previous topic? If you say yes on the previous topic, you also have to say yes here...											

Fuente: elaboración propia

Como se ha comentado en párrafos anteriores, se espera que exista una clara diferenciación entre los ítems 3.1 y 3.2, distinción que también se exige al ítem 3.3 “Los sujetos han sido asignados aleatoriamente al grupo control o experimental” por estar todos estrechamente vinculados. Sin embargo, de los tres ítems, el 3.3 es el que mejor ha sido evaluado en cuanto a su redacción y pertinencia, ambos criterios con un 90% de aprobación (ver tabla 36).

Tabla 36: Versión inicial y final del ítem 3.3

Ítem 3.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	3 10%	27 90%	1 3,6%		27 96,4%		1 3,6%	27 96,4%	1 3,3%	2 6,7%	27 90%
Versión inicial del ítem				Versión final del ítem				Referencias			
Los sujetos han sido asignados aleatoriamente al grupo control o experimental				Los sujetos han sido asignados aleatoriamente al grupo control y experimental				CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos–Álvarez y Catena (2004), Sánchez–Meca, López–Pina y Nuñez–Nuñez (2012).			
Redacción											
<ul style="list-style-type: none"><li>– Estas tres cuestiones están íntimamente relacionadas y creo que inducen a confusión.</li><li>– Redundant with the questions of category II.</li><li>– Yo colocaría este ítem de primero. Lo redactaría así: "los sujetos han sido asignados aleatoriamente a los grupos". Si hay asignación aleatoria de los sujetos a los grupos, entonces no se estaría trabajando con aulas (secciones) tal y como están conformadas naturalmente. Si se responde que no a este ítem entonces ya se sabe que los grupos son naturales y aquí la clave es que los tratamientos sean asignados aleatoriamente a los grupos.</li></ul>											

Fuente: elaboración propia

Un último ejemplo representa la integración de ítems. Tal es el caso de la *dimensión resultados* en donde se valora la interpretación y discusión de los hallazgos del estudio.

El ítem 6.1 “Describe cómo la intervención ha afectado los resultados” se sugiere fusionar con el ítem 6.2 “Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos” (ver tablas 37 y 38, respectivamente) porque son muy similares en su contenido. Por tanto, la versión final de este ítem se ha centrado en la intervención y no tanto en el informe de resultados “Se discuten las causas por las cuales la intervención ha sido efectiva o no” (ver tabla 37).

Tabla 37: Versión inicial y final del ítem 6.1

Ítem 6.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,7%	5 16,7%	23 76,7%		1 3,6%	27 96,4%		4 14,3%	24 85,7%	1 3,3%	2 6,7%	27 90%
Versión inicial del ítem			Versión final del ítem			Referencias					
Describe cómo la intervención ha afectado los resultados			Se discuten las causas por las cuales la intervención ha sido efectiva o no			CONSORT (2010), Ramos–Álvarez y Catena (2004), STARD (2003).					
Redacción											
– No me puedo imaginar un estudio de intervención en que no se comuniquen los efectos de la intervención.											
– Se describe...											
– Donde está el tópico IV Reporte de resultados, sugiere titularlo: Validez de la conclusión estadística. Hay que incluir algún ítem que haga referencia a la adecuación de los análisis realizados ¿se han comprobado supuestos del modelo estadístico? ¿se ajustan los análisis realizados a los objetivos de la investigación y a la naturaleza de los datos? Son cuestiones relativas a la validez de la conclusión estadística que pueden afectar enormemente a la calidad de la investigación.											
– Es muy parecido al siguiente ítem.											
– Buscar unificar este ítem y el siguiente porque son similares.											
– It seems very ambiguous.											
– This is vague. What do you mean with 'how'? To what degree, what underlying processes...?											
– En mi opinión, aquí es clave que en la investigación se realicen todas las comparaciones que deben realizarse en los diseños pretest–posttest, y no sólo los contrastes posttest–posttest, como sucede en muchos casos. Lo primero es que indique si la intervención tuvo efecto y luego cuál fue la dirección de ese efecto. Yo lo redactaría así: "describe cómo la intervención ha afectado al rendimiento"											

Fuente: elaboración propia

Y como se ha comentado en el párrafo anterior, el ítem 6.2 “Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos” se ha fusionado con el ítem precedente para así no duplicar información porque los comentarios cualitativos apuntan a esta misma idea (ver tabla 38).

Tabla 38: Versión inicial y final del ítem 6.2

Ítem 6.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
1 3,4%	4 13,8%	24 82,8%			27 100%	1 3,7%	2 7,4%	24 88,9%	4 13,3%	1 3,3%	25 83,3%
Versión inicial del ítem			Versión final del ítem			Referencias					
Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos			Se discuten las causas por las cuales la intervención ha sido efectiva o no			CONSORT (2010), Ramos-Álvarez y Catena (2004).					
Redacción											
– Está bien pero que algo sea fácil y sencillo es subjetivo de acuerdo a los conocimientos previos que se tengan. Esto sólo para estar pendiente de ítems como éste.											
– Yo no incluiría este ítem. ¿Qué entendemos por "claros y sencillos"? La valoración de este ítem puede diferir drásticamente en función de la base metodológica y experiencia investigadora del evaluador.											
– Se repite con el anterior.											
– Es igual que el anterior. Unificar en un solo ítem.											
– Se parece al ítem anterior.											
– You can list some minimum stats that should be reported.											
– "Informa de los efectos de la intervención de una manera fácil para su comprensión, utilizando (...)"											
Orden											
– La comunicación del estudio (reporting) no tiene por qué estar relacionada con la calidad del estudio											

Fuente: elaboración propia

Una vez elaborados todos los cambios y eliminaciones de los ítems según las valoraciones de los expertos (en el anexo 15 se pueden ver todos estos cambios con un mayor detalle) es cierto que también estas modificaciones propiciaron la creación de otros indicadores.

Un grupo nuevo de ítems se centró en aspectos vinculados a la *validez de constructo* relativos a las consideraciones teóricas, objetivos e hipótesis del estudio, entre ellos: “Se mencionan estudios previos en donde la intervención es eficaz a través de diseños aleatorios los cuales se han implementado en contextos similares” y “Los objetivos e hipótesis del estudio son claros”. Otros atendieron asuntos relacionados con la difusión/imitación del tratamiento o con la rivalidad compensatoria de los sujetos: “Se garantiza que los sujetos del grupo control y experimental no comparten información o intentan copiar el tratamiento” y “Se compensa al grupo control con técnicas cualitativas (como entrevistas no estructuradas y observación directa) para evitar que busquen competir con el grupo experimental”. Y, un último indicador focalizado en el

la medición de la VD por un solo método “La variable dependiente es medida con mas de un método” (ver tabla 39).

Tabla 39: Nuevos ítems relacionados con la validez de constructo

Nuevos ítems	Referencias
Se mencionan estudios previos en donde la intervención es eficaz a través de diseños aleatorios los cuales se han implementado en contextos similares	CONSORT (2010), Ramos-Álvarez y Catena (2004), STARD (2003).
Los objetivos e hipótesis del estudio son claros	CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Education group on guidelines on evaluation (1999), Morrison, Sullivan, Murray & Jolly (1999), Ramos-Álvarez y Catena (2004).
Se garantiza que los sujetos del grupo control y experimental no comparten información o intentan copiar el tratamiento	Cook & Campbell (1979), Cook, Campbell & Peracchio (1990)
Se compensa al grupo control con técnicas cualitativas (como entrevistas no estructuradas y observación directa) para evitar que busquen competir con el grupo experimental	Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Ramos-Álvarez y Catena (2004), Shadish, Cook & Campbell (2002)
La variable dependiente es medida con mas de un método	Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Shadish, Cook & Campbell (2002),

Fuente: elaboración propia

Por otra parte, en cuanto a la *validez interna*, la recomendación ofrecida por uno de los expertos de incluir no solo un indicador que hablara acerca de las características de la muestra sino otro relacionado con los criterios de inclusión y exclusión ha servido para la creación un nuevo ítem “Explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos”. Asimismo, los comentarios realizados en cuanto a la mortalidad global de la muestra han dado pie para la elaboración de este otro ítem “El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación” (ver tabla 40).



Tabla 40: Nuevos ítems relacionados con la validez interna

Nuevos ítems	Referencias
Explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos	Shadish, Cook & Campbell (2002), STARD (2003).
El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación	Chacón, Sanduvete y Alarcón (2005), CONSORT (2010), Ramos-Álvarez y Catena (2004), Shadish, Cook & Campbell (2002).

Fuente: elaboración propia

Y, a este listado se suman otros dos. El primero, “Un año después de finalizada la intervención, se mide nuevamente la variable dependiente” por considerar que se vincula con otros que refieren al momento de medición de la VD y la homogeneidad de los instrumentos de medida. El segundo y último, concerniente a los incentivos “Los sujetos participan en la investigación de manera voluntaria sin recibir recompensas” (ver tabla 41).

Tabla 41: Nuevos ítems relacionados con la validez interna

Nuevos ítems	Referencias
Un año después de finalizada la intervención, se mide nuevamente la variable dependiente	Chacón, Sanduvete y Alarcón (2005), Des Jarlais, Lyles, Crepaz & the TREND group (2004).
Los sujetos participan en la investigación de manera voluntaria sin recibir recompensas	Shadish, Cook & Campbell (2002)

Fuente: elaboración propia

Igualmente se crearon ítems para atender la *validez externa* en función de los comentarios de los expertos y debido a la revisión de otras referencias bibliográficas. La interpretación y generalización de resultados ha sido el norte (ver tabla 42):

Tabla 42: Nuevos ítems relacionados con la validez externa

Nuevos ítems	Referencias
Los resultados de la intervención se discuten en función de la teoría y estudios empíricos relevantes	Ramos-Álvarez y Catena (2004), STARD (2003)
A partir de los hallazgos encontrados se sugieren posibles investigaciones futuras	Ramos-Álvarez y Catena (2004)
Se generalizan los resultados de la intervención solo a la población que ha sido estudiada	Campbell & Stanley (1963), CONSORT (2010), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Ramos-Álvarez y Catena (2004).
Se generalizan los resultados de la intervención a situaciones o contextos que han sido estudiados	CONSORT (2010), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Ramos-Álvarez y Catena (2004).

Fuente: elaboración propia

Ya para finalizar este epígrafe en la tabla 43 se presentan los ítems vinculados a la *validez de conclusión estadística*, los cuales han sido sugeridos por varios expertos en las distintas valoraciones cualitativas.

Esta dimensión recoge aspectos vinculados con el tamaño de la muestra “El tamaño de la muestra tanto en el grupo control como experimental es mayor de 30 sujetos”, la fiabilidad de los instrumentos de medición “La fiabilidad de todos los instrumentos de medición es mayor a 0,80”, los supuestos del modelo estadístico “Las pruebas estadísticas aplicadas tienen en cuenta los supuestos que deben cumplir los datos”, la magnitud del efecto “Ofrece información con respecto a la magnitud del efecto” y el nivel de confianza “El nivel de confianza es de 99%”

Tabla 43: Nuevos ítems validez de conclusión estadística

Nuevos ítems	Referencias
El tamaño de la muestra tanto en el grupo control como experimental es mayor de 30 sujetos	U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003)
La fiabilidad de todos los instrumentos de medición es mayor a 0,80	Nunnally & Bernstein (1994)
Las pruebas estadísticas aplicadas tienen en cuenta los supuestos que deben cumplir los datos	Ramos-Álvarez, Moreno-Fernández, Valdés-Conroy & Catena (2008), STARD (2003).
Ofrece información con respecto a la magnitud del efecto	CONSORT (2010), Promising practices network (2013), STARD (2003).
El nivel de confianza es de 99%	Promising practices network (2013)

Fuente: elaboración propia

#### 4.4.2.5. La puntuación global

La valoración global de la escala (siendo el valor 5 la puntuación más alta) obtuvo una media de 4,6 sobre 5 puntos, dato que indica que en general los especialistas consideran que el instrumento cumple con lo esperado aunque como se ha visto, ha requerido de modificaciones.

Sin embargo, contados expertos sugieren valoraciones parciales para cada dimensión y que a la vez, el peso de éstas sean diferentes (ver tabla 44).

Tabla 44: Comentarios y sugerencias valoración puntuación global

Puntuación global
– Yo lo reduciría a tres apartados con seis valoraciones globales y tantas valoraciones parciales como ítems al final se dejen (hay que quitar los que sobran y redactar los que faltan). Se pueden dar entonces dos puntuaciones numéricas a la calidad global del estudio, pero el criterio de selección del mismo para un meta-análisis también debe considerar el peso relativo de cada uno de los apartados. – Al final, donde aparece la totalización apunta: "total de calidad..."

Fuente: elaboración propia

En la literatura científica no hay consenso con respecto a las dimensiones que deberían conformar una escala y en consecuencia, el número de ítems y su peso en la valoración global.

Autores como Campbell y Stanley (1963), Cook y Campbell (1979), Shadish, Cook y Campbell (2002) y grupos como *The Cochrane Collaboration* y *The Campbell Collaboration*, entre otros, consideran que todos los tipos de validez –de constructo, interna, externa y de conclusión estadística– deberían obtener la misma ponderación.

No obstante, otros como Farrington (2003) quien ha seguido los pasos de Pawson y Tilley (1997) sugiere que los criterios deben adaptarse a “la realidad” del estudio, a los objetivos de la investigación, a la teoría, al contexto y en especial, lo que se pretende medir. Para este autor son cinco grandes dimensiones las que conforman una escala de calidad metodológica, dimensiones que están organizadas en orden de importancia de mayor a menor peso: validez interna, validez descriptiva, validez de conclusión estadística, validez de constructo y validez externa.

Más allá de la discusión teórica entre investigadores, lo que si demuestra el estudio clásico en Medicina de Jüni, Witschi, Bloch y Egger (1999), es que los resultados clínicos que obtuvieron en su investigación dependieron de los criterios de calidad que habían sido utilizados en las escalas y no de la eficacia de los tratamientos aplicados.

Valentine y Cooper (2008) afirman que esta dificultad no solo está presente en el campo médico sino que desafortunadamente también se encuentra en las Ciencias Sociales. Sugieren que las escalas deben adaptarse a los objetivos del estudio, ser flexibles y que puedan aplicarse por personas con diferentes niveles de formación y conocimiento. Para estos autores, la suma de las diferentes dimensiones no debería ser un simple número que represente la calidad global del estudio, tal y como lo comprueba su propia escala *Study DIAD*.

Por consiguiente, después del juicio de expertos se asumió valorar independientemente cada tipo de validez con una ponderación diferente (ver anexo 16). Se asignó un 30% a la validez de constructo, validez interna y validez de conclusión estadística mientras que a la validez externa solo un 10%

information about the external validity of any single research project is the least important to a systematic reviewer since the main aims of a systematic review and meta-analysis include establishing the external validity or generalizability of results over different conditions and investigating factors that explain heterogeneity in effect size among different evaluation studies (Farrington, 2003, p. 61).

La puntuación final representó un nivel de calidad metodológica baja, media o alta. La tabla 45 compara el sistema de puntuación global de la escala, antes y después del juicio de expertos.

Tabla 45: Versión inicial y final de la puntuación global de la escala

Antes del juicio de expertos	
Valoración global de la calidad del estudio: corresponde a la suma de las puntuaciones en cada una de las siete categorías. Si el indicador ha sido valorado con un NO se puntuará con un valor de 0. Si el indicador ha sido evaluado con un SI se puntuará con un valor de 1. La valoración global de la calidad del estudio oscilará entre 0 pts y 33 pts.	
Categorías	Puntuaciones
I. Intervención	/7
II. Conformación de grupos	/6
III. Asignación	/5
IV. Muestra	/4
V. Recolección de datos	/4
VI. Reporte de resultados	/5
VII. Sesgos del investigador	/2
Total calidad global del estudio	/33

Después del juicio de expertos			
Puntuación por tipo de validez			
Tipos de validez	Valor total	%	Total
Validez de constructo	MÁX 13	30%	
Validez interna	MAX 19	30%	
Validez externa	MAX 7	10%	
Validez de conclusión estadística	MAX 5	30%	
Total	44	100%	

Calidad global estudio primario	
Niveles calidad estudios primarios	Valor
Bajo (entre 0 y 4, 99 ambos inclusive)	
Medio (entre 5 y 6,99 ambos inclusive)	
Alto (entre 7 y 10, ambos inclusive)	

Fuente: elaboración propia

#### 4.4.3. Tercera fase: fiabilidad inter-observadores

Algunos autores como Rosenthal (1984 citado por Sánchez-Meca y Ato, 1989) no recomienda el coeficiente Kappa de Cohen (1960) para estudiar la fiabilidad inter-observadores y *The Cochrane Collaboration* apunta que aunque no forma parte del procedimiento regular a seguir en sus revisiones porque podría producir resultados erróneos al existir siempre la posibilidad de desacuerdo entre mas de un revisor, si es cierto que considera que puede ser útil “para mostrar la existencia de problemas, especialmente en los estadios tempranos de la prueba piloto” (Higgins & Green, 2011, p. 170) y de esta manera, revisar los esquemas de codificación

Los expertos en un área particular frecuentemente tienen opiniones establecidas de antemano que pueden sesgar sus evaluaciones sobre la relevancia y validez de los artículos (Cooper 1989, Oxman 1993). Por lo tanto, aunque es importante que al menos un revisor tenga conocimientos sobre el área a revisar, puede ser

beneficioso tener un segundo revisor que no sea experto en el contenido (Higgins & Green, 2011, p. 169).

Por tanto, se seleccionaron al azar 15 estudios –de un total de 90– los cuales han sido valorados según su calidad metodológica por dos observadores –uno de ellos es el investigador y el otro experto en meta-análisis– (ver estudios escogidos al azar en el anexo 17).

La concordancia entre los dos observadores ha sido de 0,532 (ver anexo 18) que según la interpretación del coeficiente de Kappa se consideran un acuerdo moderado bajo los criterios establecidos por Landis y Koch (1977).

Las diferencias encontradas entre observadores permitieron mejorar aun mas algunos indicadores. La tabla 46 resume a modo de ejemplo los indicadores que bien sufrieron modificaciones en la redacción (ítems 2.4, 2.5, 9.2, 11.4, 13.2) y ajustes de contenido (ítem 1.2, 5.2, 11.3, 7.1, 7.2, 13.1) y en el anexo 19 finalmente la versión aplicada de la escala.

Tabla 46: Modificaciones de los indicadores luego de la fiabilidad inter-observadores

Ítem	Antes fiabilidad inter-observadores	Después fiabilidad inter-observadores
1.2. Estudios previos	Se mencionan estudios previos en donde la intervención es eficaz a través de diseños aleatorios los cuales se han implementado en contextos similares	Se mencionan estudios previos los cuales se han implementado en contextos similares donde la intervención ha sido eficaz
2.4. Adecuación instrumentos de medida VD	Todos los instrumentos de medición que se han aplicado para medir la variable dependiente se corresponden con su definición conceptual	Se describen todos o el único instrumento de medición que se han aplicado para medir la variable dependiente
2.5 Validación instrumentos de medida VD	Todos los instrumentos de medición que se han aplicado para medir la variable dependiente han sido validados/estandarizados	El único o todos los instrumentos de medición que se han aplicado para medir la variable dependiente han sido validados/estandarizados
5.2 Inclusión-exclusión de sujetos	Explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos	Antes de la intervención, todos los sujetos del estudio han formado parte de la investigación, no se han excluido sujetos
7.1 Mortalidad (sujetos del estudio)	El porcentaje de participantes que no ha llegado a terminar el estudio es entre 0 y 20% de la muestra original (ambos inclusive)	El porcentaje de participantes que no ha llegado a terminar el estudio es >0% y <20% de la muestra original (ambos inclusive)
7.2 Mortalidad (grupo control y experimental)	El porcentaje de la diferencia de pérdida de participantes entre el grupo control y experimental es entre el 0 y 20%	El porcentaje de la diferencia de pérdida de participantes entre el grupo control y experimental es >0% y <20%
9.2 Instrumentos de medida (homogeneidad)	Todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales	El único o todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales
11.3 Estrategia de análisis-intervención	La estrategia de análisis aplicada ha sido por intención de tratar ( <i>intention to treat analysis</i> )	La estrategia de análisis aplicada para los sujetos que han complementado o no la intervención ha sido por intención de tratar ( <i>intention to treat analysis</i> )
11.4 Limitaciones intervención	Informa de las dificultades para la implementación de la intervención	Informa de las debilidades del estudio
13.1 Tamaño muestra	El tamaño de la muestra tanto en el grupo control como experimental es entre 10 y 30 sujetos (ambos valores inclusive)	Al menos el tamaño de la muestra o bien del grupo experimental o bien del grupo control es entre 10 y 30 sujetos o mayor de 30
13.2 Fiabilidad instrumentos de medición	La fiabilidad de todos los instrumentos de medición es mayor a 0,80	La fiabilidad del único o de todos los instrumentos de medición es mayor a 0,80

Fuente: elaboración propia

Continuando con la valoración inter-observadores, solo se han eliminado los ítems 8.4 “duración del tratamiento” y 9.3 “seguimiento de la VD” por considerarse que el primero refiere a una variable moderadora ecológica que ha sido posteriormente

evaluada y el segundo, se logra o no dependiendo del objetivo de la investigación pero ambos no se relacionan directamente con la calidad metodológica de los estudios.

Estos dos cambios han afectado el porcentaje que se había asignado a los tipos de validez después del juicio de expertos, en especial, la validez interna que ahora obtiene el valor mas alto y la validez externa que también ha aumentado, modificaciones ambas que no han sido las que inicialmente se habían asumido no manteniéndose así la postura de Farrington (2003) aunque no se han perdido de vista las sugerencias de Valentine y Cooper (2008) (ver tabla 47).

Tabla 47: Tipos de validez y porcentaje asignado antes y después de la fiabilidad inter-observadores

Antes fiabilidad inter—observadores (resultados del juicio de expertos)				Después fiabilidad inter—observadores			
Puntuación por tipo de validez				Puntuación por tipo de validez			
Tipos de validez	Valor total	%	Total	Tipos de validez	Valor total	%	Total
Validez de constructo	MAX 13	30%		Validez de constructo	MAX 13	30,95%	
Validez interna	MAX 19	30%		Validez interna	MAX 17	40,47%	
Validez externa	MAX 7	10%		Validez externa	MAX 7	16,66%	
Validez de conclusión estadística	MAX 5	30%		Validez de conclusión estadística	MAX 5	11,90%	
Total	44	100%		Total	42	100%	

Fuente: elaboración propia

## 4.5. Medida de los resultados

“... while the literature retrieval and coding phases may take a large proportion of the time needed to complete a systematic review, the data analysis stage requires some careful thought about how to examine the data and understand the patterns that may exist” (Pigott, 2012, p. 13) siendo entonces necesario como paso previo al análisis de los datos, tomar decisiones relativas en cuanto a la medida de los resultados.

### 4.5.1. Programa y software estadístico

El análisis estadístico se ha realizado utilizando los softwares *Statistical Package for the Social Sciences* (SPSS, versión 17) y *Review Manager* (RevMan, versión 5.2, The Cochrane Collaboration, 2012).

Para el cálculo de la magnitud del efecto se han usado dos calculadoras disponibles online en la página web *The Campbell Collaboration* [http://www.campbellcollaboration.org/resources/effect\\_size\\_input.php](http://www.campbellcollaboration.org/resources/effect_size_input.php) Este recurso tiene implementadas las múltiples fórmulas para calcular la magnitud del efecto recogidas en libro de Lipsey y Wilson (2001) titulado *Practical Meta-Analysis*.

Asimismo, para el cálculo de la magnitud del efecto global, los intervalos de confianza y el contraste de hipótesis sobre la nulidad del efecto se utilizaron las macros de SPSS disponibles en la página web anteriormente citada.

En presencia de heterogeneidad, se ha procedido a estudiar las posibles fuentes de heterogeneidad o variables moderadoras a partir del análisis por subgrupos. Para realizar este análisis se ha utilizado también las macros de SPSS de Wilson/Lipsey.

#### **4.5.2. Unidad de análisis estadístico**

El primer paso para delimitar la unidad de análisis, es “establish the definition of the major coding unit, then specify the various levels or components of coding required for that unit. In almost all cases the primary coding unit will be a study” (Lipsey & Wilson, 2001, p. 75).

Definido qué se entiende por estudio “a set of data collected under a single research plan from a designated sample of respondents ... [y por reporte de investigación] “one written report to be described in multiple written resport” (p. 76), se atendieron los problemas de independencia a los que se enfrenta todo meta-análisis (Borenstein, Hedges, Higgins & Rothstein, 2009).

No siendo común en los meta-análisis la existencia de un único constructo de interés (Lipsey & Wilson, 2001) se ha analizado el rendimiento académico en situaciones de aprendizaje cooperativo e individual, con lo cual, se han revisado con especial atención aquellos estudios: (a) que provienen de un mismo autor o grupo de investigación o universidad(es) (independencia entre autores y equipos de investigación), (b) en donde para las mismas muestras la variable dependiente se ha medido con dos o mas medidas operativas (independencia entre medidas), (c) que incluyen un grupo control que es el mismo para la estimación de cada tamaño del efecto de mas de una intervención (independencia entre muestra) o (d) que han sido presentados por los mismos investigadores en distintos medios (independencia entre estudios) (Botella y Gambara, 2002).



De los 90 estudios que han cumplido con los criterios de inclusión se han obtenido 119 resultados (ver anexo 20). En 6 de ellos, los autores han publicado por documento mas de un proyecto de investigación o resultado (*outcome*). Cuando en un mismo documento se han presentado resultados de diferentes experimentos, estos se han considerado como estudios individuales y, de la misma forma, se ha procedido en el caso de documentos donde se mencionaban los resultados por subgrupos definidos por el semestre o año académico (ver tabla 48).

Tabla 48: Documentos que reportan mas de un proyecto de investigación o resultado

Treisman, P. (1985)	Cohort 1978-1982: 331 first year students of mathematics at University of California, Berkeley
	Cohort 1978-1983: 184 first year students of mathematics at University of California, Berkeley
McDonald, B., Larson, C., Dansereau, D. Spurlin, J. (1985)	Experiment 1: 60 students recruited from general psychology classes at Texas Christian University
	Experiment 2: 87 students recruited from general psychology classes at Texas Christian University
Ganter, S. (1994)	Fall semestre: 100 students in first semestre precalculus at the community college. University of California, Berkeley
	Spring semestre: 141 students in first semestre precalculus at the community college. University of California, Berkeley
Harding, R., Fletcher, R. (1994)	Course RSD073: 98 students in a remedial mathematics. Tenesse Technological University
	Course RSD083: 57 students in a developmental mathematics. Tenesse Technological University
Leveson, L. (1999)	Year 1: 49 students first-year Management Accounting. School of Business. La Trobe University
	Year 2: 76 students first-year Management Accounting. School of Business. La Trobe University
	Year 3: 44 students first-year Management Accounting. School of Business. La Trobe University
Nihalani, P., Mayrath, M., Robinson, D. (2011)	Experiment 1: 65 undergraduate students who majored in disciplines unlikely to require computer-networking proficiency, computer science “novices”
	Experiment 2: 62 expert undergraduate students in computer science, majors required proficiency in computer networking

Fuente: elaboración propia

En otros casos, los estudios con varios grupos experimentales y un grupo control, se han considerado tantos resultados independientes como condiciones experimentales. El número de sujetos en el grupo control fue dividido a partes iguales entre los grupos experimentales, de forma que cada sujeto se ha incluido una sola vez en el análisis. Esta situación se ha presentado en 19 documentos siendo una estrategia que también se ha

utilizado en otros meta-análisis (Acarturk, Cuijpers, Van Straten & De Graaf, 2009; Cuijpers, Van Straten, Smits & Smit, 2006; Watson & Rees, 2008) (ver tablas 49 y 50).

Tabla 49: Documentos que reportan varios grupos experimentales y un grupo control

McDonald, B., Larson, C., Dansereau, D., Spurlin, J. (1985)	Experiment 1: System (formal pair) group, Ne=20; individual study group, Nc=20* No-system (informal pair) group, Ne=20; individual study group, Nc=20*
O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J., Larson, C., Young, M. (1985)	Multiple elaboration dyads, Ne=20; single elaboration individuals, Nc=22* Single elaboration dyads, Ne=27; single elaboration individuals, Nc=22*
O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J., Young, M. (1986)	Active triads Ne=14; Individuals, Nc=14* Pasive triads Ne=16; Individuals, Nc=14* Diads Ne=17; Individuals, Nc=14*
O'Donnell, A., Larson, C., Dansereau, D., Rocklin, T. (1986)	Cooperative dyas editing, Ne=24; individual nonediting, Nc=20* Cooperative dyas nonediting, Ne=24; individual nonediting, Nc=20*
O'Donnell, A., Dansereau, D., Hythecker, V., Hall, R., Skaggs, I., Lambiotte, J., Young, M. (1988)	Prompting-only dyads, Ne=26; no-strategy individuals, Nc=18* Distributed-planning-with-prompting dyads Ne=27; no-strategy individuals, Nc=18* Preplanning-with-prompting dyads Ne=27; no-strategy individuals, Nc=18*
Oh, H. (1988)	Cooperative task, Ne=39; Control Nc=38* Cooperative incentive Ne=37; control Nc=38*
O'Donnell, A., Dansereau, D., Hall, R., Rocklin, T. (1987)	Scripted dyads, n = 30; individual group, n = 31*. Unscripted dyads, n = 32; individual group, n = 31*
Fantuzzo, J., Riggio, R., Connelly, S., Dimeff, L. (1989)	Dyadic structured format, Ne=22; Independent unstructured format, Nc=30* Dyadic unstructured format, Ne=19; Independent unstructured format, Nc=30*
Borresen, C. (1990)	Cooperative assigned groups, Ne=34, Traditional method Nc=37* Cooperative voluntary groups, Ne=35, Traditional method Nc=37*

Fuente: elaboración propia

Tabla 50: Documentos que reportan varios grupos experimentales y un grupo control (continuación)

Riggio, R., Fantuzzo, J., Connelly, S., Dimeff, L. (1991)	Dyadic structured format, Ne=20; independent unstructured format Nc=23* Dyadic unstructured format, Ne=21; independent unstructured format Nc=23*
Harding, R., Fletcher, R. (1994)	STAD-coop group, Ne=20; Non-cooperative control group, Nc=38* Ability- aligned-coop group, Ne=19; Non-cooperative control group, Nc=38*

	Alphabetical-coop group, Ne=21; Non-cooperative control group, Nc=38*
	STAD-coop group, Ne=15; Non-cooperative control group, Nc=41
Keeler, C., Steinhorst, R.(1994)	First cooperative group, Ne=41; traditional control group, Nc=51* Second cooperative group, Ne=35; traditional control group, Nc=51*
Kerr, D., Murthy, U.(1994)	GDSS-mediated, Ne=20; individual group Nc=14* Face to face group, Ne=20; individual group, Nc=14*
O'Brien, G., Peters, J.(1994)	Coop/No computers, Ne=73, No Coop/No Computers, Nc=40* Coop/Computers, Ne=30, No Coop/No Computers, Nc=40*
AbuSeileek, A. (2007)	Cooperative computer-mediated communication technique (CopCMC); Ne=36; collective traditional technique , Nc=34* Cooperative traditional technique ; Ne=36; collective traditional technique , Nc=34*
Krause, U., Stark, R., Mandl, H.(2009)	Cooperative learning without feedback, Ne=50; Individual learning without feedback, Nc=17* Cooperative learning with feedback, Ne=52; Individual learning without feedback, Nc=17*
Jalilifar, A. (2010)	Student Team Achievement Division (STAD), Ne=30; control group Nc=30* Group Investigation (GI), Ne=30; control group, Nc=30*
Koç, Y., Doymuş, K., Karaçöp, A., Şimşek, U. (2010)	Group investigation, Ne=30; control group, Nc=36* Jigsaw techniques, Ne=40; control group, Nc=36*
Sears, D., Pai, H. (2012)	Reward cooperative group, Ne=10; individual no-reward group, Nc=10* No-reward cooperative group, Ne=10; individual no-reward group, Nc=10*

Fuente: elaboración propia

Cuando en un mismo estudio se ha presentado como resultado varios *outcome* de rendimiento, bien por la aplicación de más de un instrumento de medida o bien por un mismo instrumento conformado por varias subescalas, estos se han considerado como no independientes (ver tabla 51).

Lipsey y Wilson (2001) sugieren promediar estos resultados “to get a single effect size for each construct in each study. This approach has some advantages in simplifying the coding of multiple measures but it precludes any later analysis of the influence on study findings of differing operationalizations of a construct” (p. 78).

En estos casos, un total de 36 estudios, se ha calculado para cada medida resultado de rendimiento (*outcome*) la magnitud del efecto y después se ha ponderado –y no promediado–, cada efecto no independiente por el inverso de su varianza en una única magnitud del efecto (Springer, Stanne & Donovan, 1999).

El hecho de no haber excluido algunas de estas medidas aplicando algún procedimiento aleatorio o un determinado criterio de calidad, se sustenta en el argumento de Lipsey y

Wilson (2001) quienes señalan que cuando para el análisis se dejan fuera algunas medidas se perdería la oportunidad de examinar empíricamente si pudieran ofrecer resultados diferentes.

De esos 36 estudios que presentaron varios resultados de rendimiento, en 20 se aplicaron dos medidas por *outcome*, en 5 se utilizaron 3, en 7 se emplearon 4, en 3 se administraron 5 y solo en un estudio se reportaron 6. Asimismo, se puede observar en la tabla 51 que solo 11 de esos estudios individuales comparten dos condiciones a la vez: mas de un resultado de rendimiento (*outcome*) y mas de una condición experimental.

Tabla 51: Documentos que reportan varios resultados de rendimiento

Autores	Año	Nº**
<b>*O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J., Larson, C. &amp; Young, M.</b>	1985	2
<b>*McDonald, B., Larson, C., Dansereau, D. &amp; Spurlin, J. (experimento 1)</b>	1985	2
<b>*O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. &amp; Young, M.</b>	1986	2
O'Donnell, A., Dansereau, D., Rocklin, T., Larson, C., Hythecker, V., Young, M. & Lambiotte, J.	1987	2
<b>*O'Donnell, A., Dansereau, D., Hall, R. &amp; Rocklin, T.</b>	1987	2
Jimison, L.	1990	2
Dees, R.	1991	2
Courtney, D., Courtney, M. & Nicholson, C.	1992	2
Lambiotte, J., Skaggs, L. & Dansereau, D.	1993	2
Hwong, N., Caswell, A., Johnson, D. & Johnson, R.	1993	2
Ganter, S.	1994	2
<b>*O'Brien, G. &amp; Peters, J.</b>	1994	2
<b>*Kerr, D. &amp; Murthy, U.</b>	1994	2
Giraud, G.	1997	2
Feyzioğlu, B., Akçay, H. & Şahin-Pekmez, E.	2007	2
<b>*AbuSeileek, A.</b>	2007	2
<b>*Koç, Y., Doymuş, K., Karaçöp, A. &amp; Şimşek, U.</b>	2010	2
Bayraktar, G.	2011	2
Nihalani, P., Mayrath, M. & Robinson, D.	2011	2
Lin, Z.	2012	2
Clifford, J.	1981	3
<b>*Oh, H.</b>	1988	3
Turner, J.	1995	3
Yaeger, P., Marra, R., Costanzo, F. & Gray, G	1999	3
Doymuş, K., Şimşek, U. & Karaçöp, A.	2009	3
<b>*O'Donnell, A., Larson, C., Dansereau, D. &amp; Rocklin, T.</b>	1986	4
Lambiotte, J., Dansereau, D., Rocklin, T., Fletcher, B., Hythecker, V., Larson, C. & O'Donnell, A.	1987	4
Iwasiw, C. & Goldenberg, D.	1993	4
George, P.	1994	4
Doymuş, K.	2008a	4
<b>*Sears, D. &amp; Pai, H.</b>	2012	4
Hsiung, C.	2012	4
Basili, P.	1988	5
Doymuş, K.	2007	5
Cheng, K.	2009	5
Rewey, K., Dansereau, D., Skaggs, L., Hall, R. & Pitre, U.	1989	6

\* corresponde a los estudios individuales que comparten a la vez la condición de mas de un *outcome* de rendimiento y condición experimental

\*\* número de *outcomes* de rendimiento

Fuente: elaboración propia

#### 4.5.3. Magnitud del efecto global

En el meta-análisis, es necesario sintetizar los resultados obtenidos de distintas investigaciones en una métrica en común, es decir, en términos comparables, lo que permite no solo “hacer valoraciones globales de los resultados, sino también (...) detectar regularidades, que se expresarán como covariaciones entre la magnitud del resultado y las características de los estudios” (Botella y Gambara, 2002, p. 57).

Al respecto, *effect size* en inglés y en español magnitud del efecto representa “... la eficacia cuantificada entre los distintos niveles de la variable independiente, complementado así la información ofrecida por la probabilidad de ocurrencia de la hipótesis nula, pues además de confirmar la existencia de diferencias también informa

de la magnitud de éstas” (Tejero–González, Castro–Morera y Balsalobre–Fernández, 2012, p. 717). Indica el grado en que un fenómeno está presente en la población “Effect sizes calculated to describe the data in a sample, like any other descriptive statistic, also potentially estimate the corresponding population parameter” (Fritz, Morris & Richler, 2012, p. 3).

Es un valor distinto a cero de la población y muestra el grado en que la hipótesis nula es falsa. La magnitud del efecto se ha cuantificado según Hedges y Olkin (1985) –que tiene un carácter inferencial cuando se compara con la fórmula de Cohen (1988) que tiene un carácter mas descriptivo–, usando la diferencia media estandarizada  $d$  entre grupo experimental y grupo control. Se ha calculado siempre de forma que un efecto a favor del grupo experimental sea positivo:

$$d = \frac{\bar{X}_e - \bar{X}_c}{\sqrt{\frac{(n_e - 1)S_e^2 + (n_c - 1)S_c^2}{n_e + n_c - 2}}} \quad (\text{Ec.1})$$

Donde:

$\bar{X}_e, \bar{X}_c$  es la media en el grupo experimental y control respectivamente

$n_e, n_c$  es el tamaño muestral para el grupo experimental y control respectivamente

$S_e^2, S_c^2$  es la varianza muestral para el grupo experimental y control respectivamente

Y, la desviación estandarizada para la magnitud del efecto se ha calculado como:

$$se = \sqrt{\frac{n_e + n_c}{n_e n_c} + \frac{d^2}{2(n_e + n_c)}} \quad (\text{Ec. 2})$$

En los estudios con diseño pre–post, la magnitud del efecto se ha calculado a partir del cambio medio pre y post tratamiento (*gains*); así, se aumenta la precisión al controlar el efecto por las diferencias pre–tratamiento entre grupos de las variables dependientes.

Cuando el cambio medio no estaba presente, la magnitud del efecto se ha calculado a partir de las medidas post–tratamiento, como resultado del análisis sin ajustar. En

algunos estudios, solo se daban resultados ajustados, que son los que se han utilizado para el cálculo.

Cuando la información para el cálculo de  $d$  no estaba disponible, se han utilizado otros datos estadísticos como el valor del  $t$ -test, el valor  $F$  del ANOVA o  $p$ -valores. También en contados casos ha sido posible calcular el valor de  $d$  cuando el resultado era una variable de naturaleza cualitativa.

Para tamaños muestrales pequeños –a partir de 10– se ha aplicado la corrección de Hedges y Olkin, multiplicando cada efecto  $d$  por el factor de corrección  $c$  “porque esta corrección es virtualmente indiferente con muestras grandes, pero con muestras pequeñas si tiene consecuencias apreciables” (Botella y Gambara, 2002, p. 71):

$$c = 1 - \frac{3}{4(n_e + n_c) - 9} \quad (\text{Ec. 3})$$

Una vez calculado el índice  $d$  como estimador de la magnitud del efecto se valoró su tamaño. Este índice se interpreta como a cuántas desviaciones estándar se encuentra la media del grupo experimental frente a la medida del grupo control.

Como muy bien recomienda Cohen (1977, 1988 citado por Lipsey & Wilson, 2001) su valoración debe ofrecerse en el contexto de la pregunta de investigación. Sugiere una magnitud del efecto grande cuando el tamaño es  $\geq 0,8$ , moderado igual a 0,5 y pequeño  $\leq 0,2$ .

Autores clásicos como Glass, McGaw & Smith (1981) consideran que aplicar este índice en determinadas áreas de investigación de una manera “ciega” es inapropiado. Una interpretación correcta es comparar los resultados obtenidos en el nuevo meta-análisis con resultados previos en la literatura (Thomson, 2008 citado por Sun, 2008).

Otros como Botella y Gambara (2002) apuntan que en Psicología y Educación es frecuente obtener magnitudes del efecto entre 0,5 y 1. Mientras que Valentine y Cooper (2003) hacen alusión a lo que ocurre específicamente en el campo educativo “Because some areas, like education, are likely to have smaller effect sizes than others, using Cohen’s labels may be misleading” (p. 5).

#### **4.5.4. Estimación global del efecto**

Existe un gran consenso en que la precisión en la estimación de la magnitud del efecto se relaciona directamente con el tamaño muestral del estudio (Sánchez-Meca, Marín y Huedo, 2006) pero que para el cálculo de esa estimación es necesario tomar una decisión en cuanto a qué modelo estadístico asumir: el modelo de efectos fijos o el modelo de efectos aleatorios

Under the fixed-effect model we assume that there is one true effect size that underlies all the studies in the analysis, and that all differences in observed effects are due to sampling error (...) under the random-effects model we allow the true effect sizes to differ—it is possible that all studies share a common effect size, but it is also possible that the effect size varies from study to study (Borenstein, Hedges, Higgins & Rothstein, 2010, p. 97).

Por consiguiente, el método elegido para combinar los resultados de los diferentes estudios en una medida global del efecto vendrá determinado fundamentalmente por los resultados derivados del análisis de heterogeneidad.

El estimador del efecto combinado se ha calculado como una media ponderada de los estimadores de cada estudio, usando como pesos  $w$  los inversos de la varianza para cada magnitud del efecto independiente. De esta forma, efectos más precisos (derivados de estudios con mayor  $n$ ) aportan mas peso

$$w = \frac{1}{se^2} \quad (\text{Ec. 4})$$

En un primer momento, la estimación global del efecto se ha calculado bajo el modelo de efectos fijos y el modelo de efectos aleatorios pero visto el análisis previo de homogeneidad, se ha utilizado un modelo de efectos fijos para el análisis de los subgrupos porque este modelo “asume que existe un único efecto en la población y no tiene en cuenta la variabilidad de los resultados entre los distintos estudios. Así, el tamaño del estudio y su propia varianza (variabilidad intra-estudio) son los únicos determinantes de su peso en el meta-análisis” (Xunta de Galicia, sin fecha, p. 13).

Una vez obtenida la magnitud del efecto global, el siguiente paso ha sido dar respuesta a si esa magnitud del efecto global es estadísticamente significativa, es decir, si el grupo experimental (aprendizaje cooperativo) presenta resultados estadísticamente diferentes del grupo control (aprendizaje individual) siendo necesario para ello establecer el



intervalo de confianza. Este indica el intervalo dentro del cual la media de la población es probable que se encuentre a partir de los datos observados siendo muy útil para estimar con precisión la media de la magnitud del efecto y si, adicionalmente ese intervalo de confianza no incluye al cero, la media de la magnitud del efecto es estadísticamente significativa al nivel especificado por ese intervalo de confianza (Borenstein, Hedges & Rothstein, 2007).

Se ha calculado la estimación global del efecto con un intervalo de confianza de un 95%, es decir, la hipótesis nula ha sido rechazada a nivel de confianza 0,05 sin olvidar que los límites superior e inferior del intervalo de confianza así como su amplitud “informan del grado de precisión de la estimación del efecto medio de la población ... [y que como se ha asumido un modelo de efectos fijos] sólo podremos generalizar los resultados del meta-análisis a una población de estudios de características idénticas a las de los estudios incluidos” (Sánchez– Meca, Marín y Huedo, 2006, p. 13–14).

Los resultados se han representado en una gráfica llamada *forest plot* que muestra las estimaciones del efecto individual de cada estudio, además de la magnitud del efecto global y su correspondiente intervalo de confianza.

#### **4.5.5. Heterogeneidad de los estudios**

“Cuantificar la heterogeneidad en un grupo de estudios es uno de los objetivos del meta-análisis a la vez que uno de sus aspectos mas problemáticos. Su importancia deriva de que buena parte de la decisión de asumir uno u otro modelo estadístico [modelo fijo o modelo aleatorio], va a depender de esa evaluación” (Huedo–Medina y Johnson, 2010, p. 20).

En el modelo de efectos fijos la única variabilidad asumida en los estudios individuales es la debida al error del muestreo aleatorio que se cuantifica estimando la varianza intra–estudio. Esto implica que las diferencias de las magnitudes del efecto entre los estudios individuales se deben tan solo al hecho de que estas investigaciones utilizan muestras de sujetos diferentes. Requiere ponderar cada estimación de la magnitud del efecto por la inversa de la varianza y, el factor de ponderación de cada índice viene determinado exclusivamente por la varianza intraestudio.

Por tanto, la generalización de resultados se limita a la población de estudios de características similares a los incluidos en el meta-análisis “one major consideration in

the choice to use fixed-effects analyses is the nature of the generalizations desired (Konstantopoulos & Hedges, 2004, p. 293).

En este modelo, la hipótesis de homogeneidad entre las magnitudes del efecto se contrasta planteado como hipótesis nula que los efectos de los estudios son idénticos, son los mismos (Sánchez-Meca, Marín y Huedo, 2006).

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k = \theta$$

Para contrastar la hipótesis nula de homogeneidad, es decir, la evaluación del grado de heterogeneidad de los estudios, se ha empleado el estadístico  $Q$  propuesto por DerSimonian y Laird (1986)

Si la hipótesis de homogeneidad no se rechaza, entonces puede concluirse que los tamaños del efecto individuales no son mas heterogéneos entre sí de lo que cabría esperar por el azar o, lo que es lo mismo, por el error de muestreo ... [y si se rechaza] los tamaños del efecto de los estudios no son homogéneos entre si, por lo que deberían existir características diferenciales entre los estudios que deben estar provocando tal heterogeneidad (Sánchez-Meca, Marín y Huedo, 2006, p. 196).

Son precisamente esas características diferenciales las que se llaman variables moderadoras.

Si el número de estudios es pequeño, la capacidad del estadístico  $Q$  para detectar heterogeneidad es muy baja (poca potencia de contraste), mientras que, por el contrario, cuando el meta-análisis combina gran número de estudios, el resultado puede ser estadísticamente significativo incluso cuando la magnitud de la heterogeneidad no sea de relevancia. Por lo que, también se ha calculado el índice  $I^2$ , propuesto por Higgins y Thompson (2002) que mide el grado de inconsistencia entre los resultados de los diferentes estudios, expresado como la proporción de la variación entre los estudios respecto de la variación total aunque algunos autores consideran que el índice  $I^2$  es complementario del estadístico  $Q$  pero mantiene el mismo problema de potencia con un pequeño número de estudios (Huedo-Medina, Sánchez-Meca, Marín & Botella, 2006).

Si el estadístico  $Q$  informa de la existencia o no heterogeneidad en los estudios, el índice  $I^2$  permite estimar el grado de esa heterogeneidad expresada en términos porcentuales: la proporción de la variación total que es atribuible a la heterogeneidad (Higgins & Thompson, 2002; Higgins, Thompson, Deeks & Altman, 2003). Algunos autores proponen que hasta un 30% se habla de baja heterogeneidad (buena

consistencia), entre el 30% y el 50% se dice que la heterogeneidad es moderada y por encima de 75% se califica de alta.

#### 4.5.6. Variables moderadoras

La idea del análisis de subgrupos o estratificación consiste en efectuar el meta-análisis en diferentes grupos de estudios, reunidos según características de los mismos (características contextuales, metodológicas o ecológicas) y comprobar cómo varía el resultado al estimarlo de nuevo en cada ocasión.

Independientemente del modelo que se asuma, en los casos en los que exista una alta heterogeneidad entre estudios es necesario estudiar las posibles fuentes asociadas a esa variabilidad mediante la búsqueda de variables moderadoras. Esta fase ofrece la posibilidad de “... dar explicación a contradicciones o efectos que no es posible analizar con tan solo un estudio experimental, dado que cuenta con la posibilidad de agrupar más de un estudio con características similares” ofreciendo así aclaraciones a aquello que resulta contradictorio (Huedo–Medina y Johnson, 2010, p. 25).

Para estudiar y comparar los efectos de las covariantes de interés se ha utilizado un método análogo al análisis de la varianza ANOVA, que consiste en descomponer el estadístico  $Q$ , que representa la variabilidad total, en la variabilidad entre grupos  $Q_b$  (*between groups*) y  $Q_w$  la variabilidad intragrupos (*within groups*) (Konstantopoulos & Hedges, 2004).

La relación entre el estadístico  $Q_b$  y  $Q_w$  es complementaria. Por ejemplo, si la prueba  $Q_b$  es significativa, es decir, existen diferencias entre grupos y la prueba  $Q_w$  no resulta significativa, es decir, no hay variabilidad intragrupos, se puede afirmar que la variable moderadora por la que se ha agrupado a los estudios, explica un alto porcentaje de la heterogeneidad de la magnitud del efecto (Huedo–Medina y Johnson, 2010).

El estadístico  $Q_b$  se utiliza para contrastar la hipótesis de tamaños del efecto medio de los subgrupos considerados homogéneos. Si el estadístico  $Q_w$  resulta ser estadísticamente significativo, entonces existen otras variables moderadoras afectando a la variabilidad de los tamaños del efecto.

El análisis de las variables moderadoras ha sido preespecificado porque se ha determinado antes de la recolección de los resultados y su justificación científica proviene de otras fuentes diferentes a los estudios primarios incluidos (Bowen, 2000;

Johnson & Johnson, 1989; Johnson & Johnson, 2002a; Johnson, Johnson & Stanne, 2000; Johnson, Maruyama, Johnson, Nelson & Skon, 1981; Kyndt et al. 2013; Lou, Abrami & D'Apollonia, 2001; Lou, Abrami & Spence, 2000; Lou, Abrami, Spence, Poulsen, Chambers & D'Apollonia, 1996; Qin, Johnson & Johnson, 1995; Springer, Stanne & Donovan, 1999; Stanne, Johnson & Johnson, 1999)

La selección de las características (...) debe estar (...) apoyada por pruebas provenientes de fuentes diferentes a los estudios incluidos (...) De ser posible, los revisores deben preespecificar en el protocolo las características que posteriormente se incluirán en los análisis de subgrupos (...) Al preespecificar las características se reduce la probabilidad de hallazgos espurios, en primer lugar porque limita el número de subgrupos investigados y en segundo lugar porque evita que el conocimiento de los resultados de los estudios influya en qué subgrupos se analicen (Centro Cochrane Iberoamericano, 2012, p. 294).

#### **4.5.7. Análisis de sensibilidad**

El análisis de sensibilidad responde a la pregunta “What happens if some aspect of the data or the analysis is changed?” (Greenhouse & Iyengar, 2004) lo que ha permitido estudiar la influencia individual de cada estudio en el resultado del meta-análisis y, por lo tanto, determinar si los hallazgos pueden verse sesgados por estudios con escasa calidad metodológica o trabajos no publicados, entre otras razones

Cuando los análisis de sensibilidad muestran que el resultado y las conclusiones generales no se ven afectados por las diferentes decisiones que se pudieron tomar durante el proceso de revisión, los resultados de una revisión se pueden considerar como de alto grado de certidumbre (Centro Cochrane Iberoamericano, 2012, pp. 298–299).

Consiste en replicar el meta-análisis quitando en cada paso uno de los estudios incluidos, para ver si se obtienen o no resultados similares de forma global pero hay que matizar porque “...se trata de una técnica heurística y como tal debe ser interpretada con cautela, ya que no planteamos una hipótesis de investigación que luego se verifica con los datos, sino que, revisando los datos, proponemos hipótesis que los explican” (Molinero, 2008, p. 115).

#### 4.5.8. Sesgo de publicación

Los sesgos de informe

surgen cuando la diseminación de los hallazgos de las investigaciones está influenciada por la naturaleza y la dirección de los resultados. Es más probable que se publiquen los resultados “positivos” estadísticamente significativos que indican que una intervención funciona, más probable que se publiquen más rápido, más probable que se publiquen en inglés, más probable que se publiquen más de una vez, más probable que se publiquen en revistas de gran impacto y, relacionado con este último punto, y más probable que se citen por otros (Higgins & Green, 2011, p. 308).

El sesgo de publicación, entendido como la publicación selectiva de estudios sobre la base de sus hallazgos, se valora mediante métodos gráficos y analíticos.

Como método gráfico se ha utilizado el *funnel plot*, en el que se representa el tamaño muestral de cada trabajo frente al tamaño del efecto detectado “The expectation is that should appear symmetric with respect to the distribution of effect sizes and funnel shaped if no bias is present” (Sutton, 2004, p. 437).

En cuanto a los métodos analíticos, se ha aplicado la prueba de Egger para contrastar la hipótesis nula de ausencia de sesgo de publicación. Esta prueba detecta asimetría en el gráfico de embudo, sugestiva de la presencia de sesgo de publicación. El método consiste en ajustar una recta de regresión a los puntos con abscisa igual a la precisión, que es el inverso del error estándar (variable independiente), y ordenada igual al efecto estandarizado (variable dependiente). La regresión se pondera por el inverso de la varianza

Si no hay sesgo de publicación, la ordenada en el origen será igual a cero; cuanto más alejada de cero, más pronunciada es la asimetría, y mayor la evidencia de sesgo. La prueba de Egger está basada en la prueba de significación para el término independiente de la recta ajustada (Xunta de Galicia, sin fecha, p. 16).

El sesgo de publicación contra los resultados estadísticamente no significativos es un problema que se produce por la reconocida renuencia de los editores de las revistas científicas a aceptar estudios con hallazgos estadísticamente no significativos por considerarlos de menor interés. Esto se traduce en que la proporción de estudios con resultados significativos en las revistas, sea mayor que si esa proporción se calculase sobre el total de estudios realizados (Hopewell, Loudon, Clarke, Oxman & Dickersin, 2009).

En este sentido, se han descrito varios índices que permiten responder a la pregunta de cuántos estudios con magnitud del efecto medio igual a 0 serían necesarios para que al añadirlos a los  $k$  estudios recuperados la magnitud del efecto global se redujese hasta una cantidad especificada.

En este meta-análisis se ha utilizado el Número de seguridad de Orwin. Para su cálculo es necesario conocer el número de estudios ( $k_r$ ), la magnitud del efecto medio para los estudios recuperados ( $d_r$ ), una magnitud del efecto medio para los estudios no recuperados ( $d_a$ ), y una magnitud del efecto de criterio ( $d_c$ ), que puede variar según el tema de investigación. Según Cohen, la magnitud del efecto criterio de 0,2 puede interpretarse como pequeño, de 0,5 como medio y de 0,8 como grande. La fórmula para calcular el número de Orwin es la siguiente:

$$N_s = \frac{k_r(\overline{d_r} - \overline{d_c})}{d_c - \overline{d_a}} \quad (\text{Ec. 5})$$

Si este número es muy elevado, se considera que la probabilidad de que el sesgo de publicación haya modificado sustancialmente los resultados es baja.

$$Q = \sum w_i(T_i - \bar{T})^2 \quad (\text{Ec. 6})$$



## **Capítulo V. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS**

La presentación de los resultados de un meta-análisis la conforma una descripción general de la muestra así como una síntesis cuantitativa y cualitativa como producto de la integración y comparación de los estudios incluidos. Los índices estadísticos globales de la media de cada estudio y de la media de todos se acompañan del índice de homogeneidad y se organizan en función de posibles variables moderadoras del efecto (Sánchez-Meca y Ato, 1989).

Es en este capítulo donde se exhiben las características de los estudios, la magnitud global del efecto y de cada uno de esos estudios integrados, si cambian o no de dirección y de significación estadística y, en caso de confirmarse heterogeneidad ofrecer una explicación en función de sus características diferenciales siendo la calidad metodológica otra posible variable explicativa. La valoración mediante gráficos y métodos analíticos de la publicación selectiva de estudios sobre la base de sus hallazgos también forma parte de los resultados analizados.

La interpretación guiada por los objetivos del presente trabajo han permitido una discusión teórica de la eficacia del aprendizaje cooperativo frente al individual en el rendimiento académico de estudiantes y de las condiciones mas eficaces para su puesta en práctica en la universidad.

### ***5.1. Resultados del análisis***

#### **5.1.1. Descripción de la muestra**

De un total de 3744 estudios que aparecieron registrados bajo los descriptores *Cooperative Learning* y *Collaborative Learning* en las fuentes primarias y secundarias seleccionadas, se han encontrado 2536, de los cuales 2446 han sido excluidos y solo 90 han cumplido con los criterios de inclusión (ver tabla 52).



Tabla 52: Número total de estudios provenientes de la revisión de la literatura

Revisión de la literatura	Nº de estudios			
	Estudios registrados	Estudios potenciales encontrados	Estudios excluidos	Estudios incluidos
Fuentes primarias	2960	1980	1927	53
Fuentes secundarias	784	556	519	37
Total	3744	2536	2446	90

Fuente: elaboración propia

Por tanto, el número total de estudios han sido 90 documentos que reportan 119 resultados (*outcomes*) en donde han participado 8114 estudiantes universitarios (ver tabla 53 y en el anexo 22 un resumen que ofrece una descripción detallada de las características de la muestra).

Tabla 53: Muestra final

				Percentil		
Muestra final	Media	Mediana	Desviación estándar	25	50	75
8114	91,78	77	81,42	54,25	77	105,75

Fuente: elaboración final

#### 5.1.1.1. Variables extrínsecas

La tabla 54 resume las características principales de aspectos vinculados a la identificación y publicación del estudio como año de publicación, tipos de fuente, tipos de documento y lugar de procedencia de las instituciones.

Tabla 54: Resumen características extrínsecas de la muestra

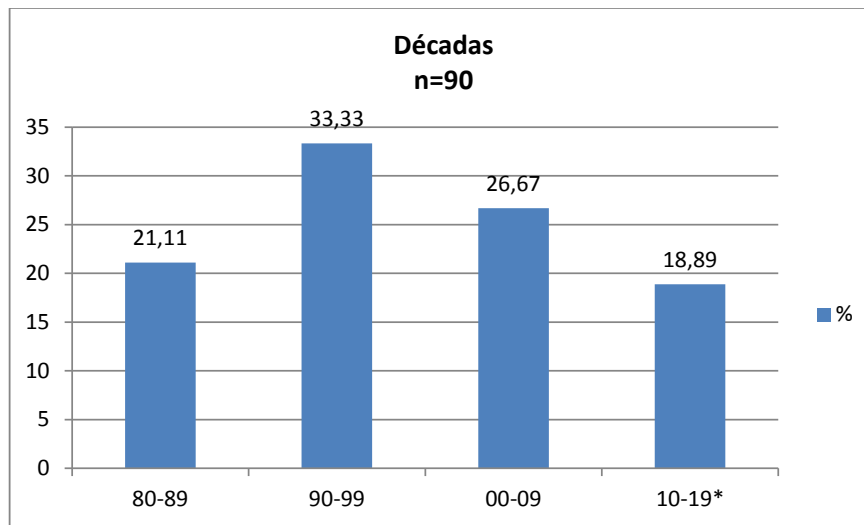
<b>VARIABLES MODERADORAS EXTRÍNSECAS</b>		
<b>Años de publicación</b>	<b>N</b>	<b>%</b>
80–89	19	21,11%
90–99	30	33,33%
00–09	24	26,67%
10–19*	17	18,89%
<b>Tipos de fuente</b>	<b>N</b>	<b>%</b>
Fuente primaria	53	58,89%
Meta-análisis	29	32,22%
Otro estudio	7	7,78%
Revisión narrativa	1	1,11%
<b>Tipos de documento</b>	<b>N</b>	<b>%</b>
Artículos de investigación	68	75,56%
Disertaciones/tesis	15	16,67%
Reporte técnico	7	7,78%
No publicados	0	0%
<b>Lugar de procedencia</b>	<b>N</b>	<b>%</b>
América	61	67,78%
Europa	17	18,89%
Asia	9	10,00%
Asia y América	1	1,11%
África	1	1,11%
Oceanía	1	1,11%

\* los criterios de inclusión señalan estudios hasta el 2012 inclusive

Fuente: elaboración propia

Se observa en el gráfico 1 como el mayor número de estudios seleccionados se encuentran en la *década* de los 90 con un 33,33% seguido por un 26,67% entre los años 2000–2009, un 21,11% entre los años 1980–1989 y con un porcentaje menor, un 18,89% entre los años 2010–2019 pero con la diferencia que este último dato corresponde solo a dos años –desde el 2010 hasta el 2012 inclusive– por tratarse del límite superior establecido para la búsqueda de documentos.

Gráfico 1: Resultados por décadas estudios incluidos

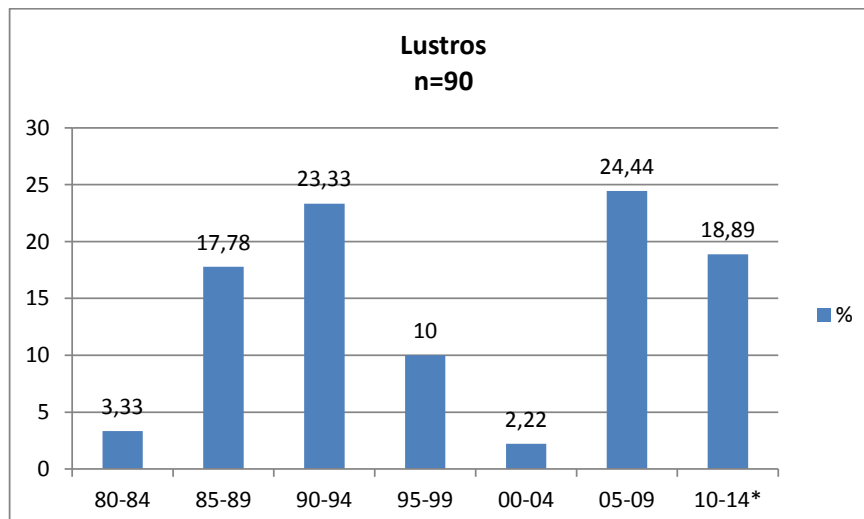


\* los criterios de inclusión señalan estudios hasta el 2012 inclusive

Fuente: elaboración propia

Un análisis más detallado indica que el mayor número de estudios por *lustros* que cumplen con los criterios de inclusión se concentran en orden de mayor a menor producción entre los años 2005–2009 (24,44%), 1990–1994 (23,33%), 2010–2014 (18,89%) –recordemos solo desde el 2010 hasta el 2012 inclusive– y 1985–1989 (17,78%). En contraparte los tres lustros que reportaron un menor número de estudios en orden decreciente son los años 1995–1999, 1980–1984, y 2000–2004 (ver gráfico 2).

Gráfico 2: Resultados por lustros estudios incluidos

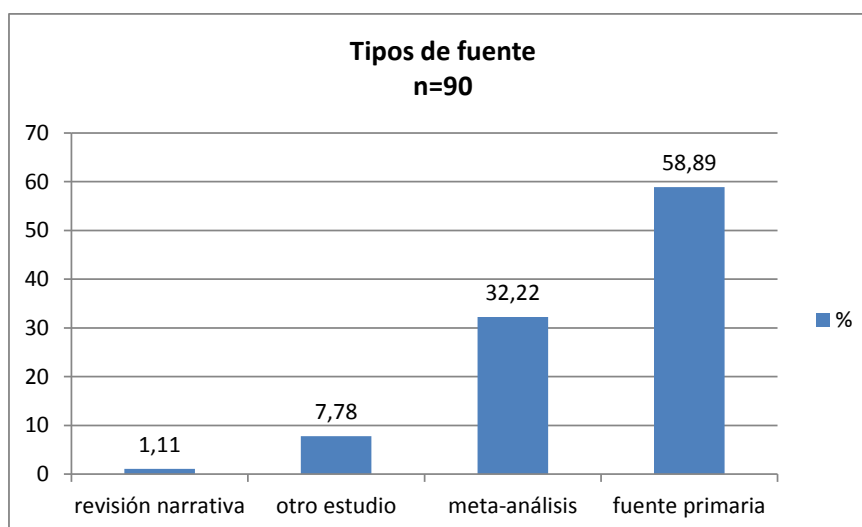


\* los criterios de inclusión señalan estudios hasta el 2012 inclusive

Fuente: elaboración propia

Los *tipos de fuente* de donde provienen más de la mitad de los estudios primarios, un 58,89%, corresponden a fuentes primarias mientras que un 41,11% a fuentes secundarias. Específicamente, un 32,22% de esas revisiones secundarias procede de otros meta-análisis, un 7,78% del repaso de la bibliografía a partir de otros estudios y un 1,1% de la revisión narrativa (ver gráfico 3).

Gráfico 3: Resultados por tipos de fuente estudios incluidos

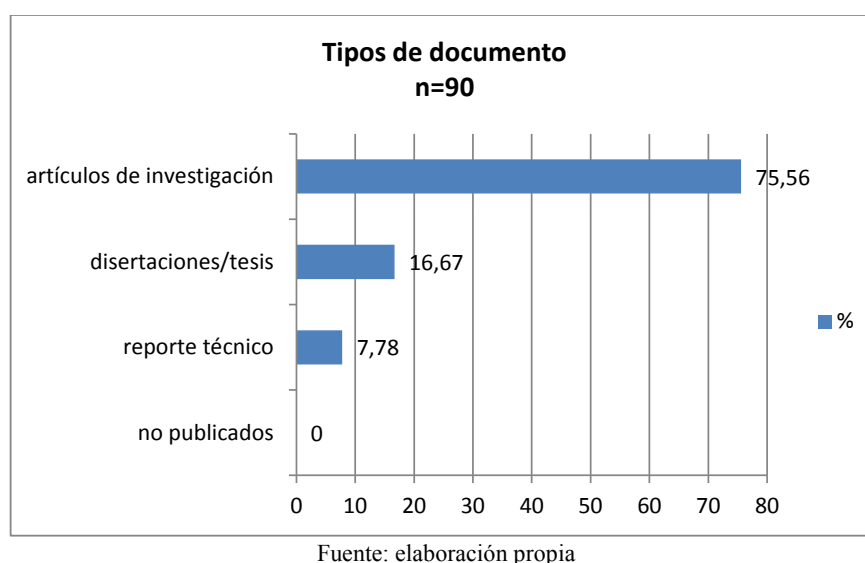


Nota: las fuentes secundarias están conformadas por meta-análisis, otros estudios y una revisión narrativa

Fuente: elaboración propia

A su vez, los *tipos de documento* que predominan son las publicaciones en revistas científicas (75,56%). Las disertaciones y tesis doctorales (16,67%) ocupan un segundo lugar muy distante cuando se comparan con el primero y, las publicaciones encontradas a través de reportes técnicos no suman más del 7,78% de los estudios primarios. La ausencia de documentos no publicados se debe a que no se obtuvo respuesta por parte de autores y grupos de investigación ante la solicitud de estos documentos (ver gráfico 4).

Gráfico 4: Resultados por tipos de documento de los estudios incluidos



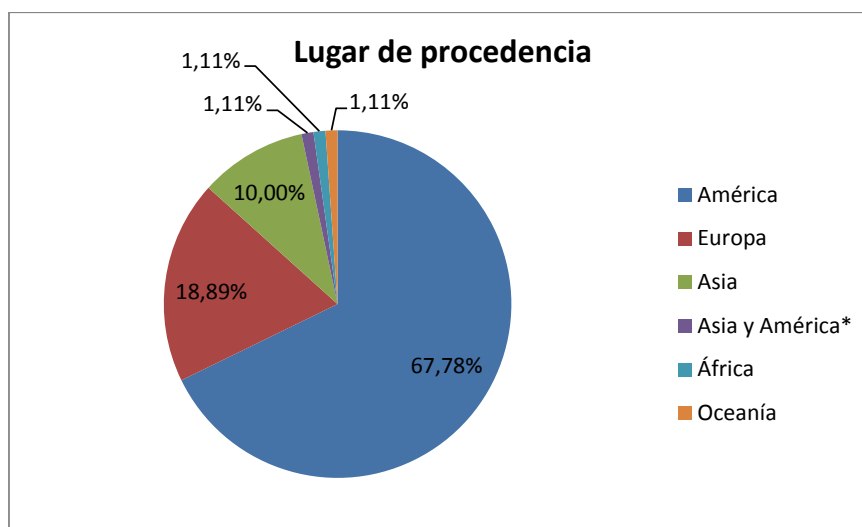
El *lugar de procedencia* en donde se han llevado a cabo las investigaciones son en su mayoría en América, con un 67,78%, dato que es representativo solo de Norteamérica específicamente de Estados Unidos, en donde se ha estudiado con mayor sistematicidad y durante el mayor período de tiempo esta metodología.

Europa ocupa el segundo lugar con 18,89%. Llama la atención que de los 17 estudios europeos, uno solo proviene de una universidad en Alemania (Krause, Stark & Mandl, 2009) y los 16 restantes pertenecen todos a universidades en Turquía con una producción reciente entre los años 2007–2011 (Bayraktar, 2011; Bilgin, 2009; Bölükbaş, Keskin & Polat, 2011; Doymuş, 2007, 2008a; Doymuş, Karacop & Simsek, 2010; Doymuş, Şimşek & Karaçöp, 2009; Feyzioğlu, Akçay & Şahin–Pekmez, 2007; Gömleksiz, 2007; Karababa, 2009; Kiliç, 2008; Koç, Doymuş, Karaçöp & Şimşek, 2010; Maden, 2010; Perihan & Kamuran, 2007; Şahin, 2010; Tanel & Erol, 2008).

El tercer lugar, pertenece a Asia con un 10% mientras que el continente africano y Oceanía son los que reportan el menor número de estudios cada uno con 1,11%.

En un solo caso, han participado en un mismo estudio mas de un continente (1,11%) dato que indica la baja participación entre universidades de distintos hemisferios (ver gráfico 5).

Gráfico 5: Resultados lugar de procedencia estudios incluidos



\* estudio de Cheng & Ku (2009)

Fuente: elaboración propia

Las revistas científicas en donde se han publicado los artículos de investigación abarcan distintos campos como Educación, Psicología, Tecnología, Ingeniería, Ciencias de la Salud, Ciencias Naturales y Matemáticas, entre otros.

De las 53 revistas, en 9 de ellas se han aceptado mas de un estudio primario seleccionado: *Journal of Educational Psychology* (7), *Contemporary Educational Psychology* (3), *The Journal of Social Psychology* (3), *Educational Research and Reviews* (3), *Journal of Chemical Education* (2), *Journal of Experimental Education* (2), *Journal of Research on Computing in Education* (2), *Learning and Instruction* (2) y *Nurse Education Today* (2) (ver tabla 55)

Tabla 55: Listado revistas científicas estudios incluidos

Revistas científicas*	
1. Applied Cognitive Psychology	28. Journal of Reading
2. Asia-Pacific Journal of Teacher Education	29. Journal of Research in Mathematics Education
3. Au Journal	<b>30. Journal of Research on Computing in Education</b>
4. Biochemical Education	31. Journal of Social Behavior and Personality
5. College Teaching	32. Journal of Statistics Education
6. Computer Assisted Language Learning	33. Journal of Turkish Science Education
7. Computers in Human Behavior	34. Journal Social Science
<b>8. Contemporary Educational Psychology</b>	35. Jurnal Sains Dan Matematik
9. Educational Research and Evaluation	36. Latin-American Journal of Physisc Education
<b>10. Educational Research and Reviews</b>	<b>37. Learning and Instruction</b>
11. Educational Technology Research and Development	38. Medical Teacher
12. Educational Techonology & Society	<b>39. Nurse Education Today</b>
13. Eurasian Journal of Educational Research	40. Primus
14. European Journal of Engineering Education	41. Research in Higher Education Journal
15. Focus On Learning Problems in Mathematics	42. Research in Science Education
16. H.Ü. Eğitim Fakültesi Dergisi	43. Research in The Teaching Of English
17. Higher Education Research & Development	44. Scientific Research and Essay
18. International Education Studies	45. System
19. Journal of Advance Nursing	46. Teaching Statistics
<b>20. Journal of Chemical Education</b>	47. The Journal of Abnormal and Social Psychology
<b>21. Journal of Educational Psychology</b>	48. The Journal of Economic Education
22. Journal of Elementary Science Education	49. The Journal of Experimental Education
23. Journal of Engineering Education	<b>50. The Journal of Social Psychology</b>
24. Journal of Excellence In College Teaching	51. The Turkish Online Journal of Educational Techonology
<b>25. Journal of Experimental Education</b>	52. World Applied Sciences Journal
26. Journal of Information Systems	53. Written Communication
27. Journal of Personality And Social Psychology	

\* las revistas se han ordenado por orden alfabético

Fuente: elaboración propia

Los *grupos de investigación* con Crooks como investigador principal han publicado todos sus estudios en la revista *Journal of Research on Computing in Education* (2) y otros en mas de una, como es el caso de Dansereau y colaboradores quienes se han decantado por revistas como *Journal of Educational Psychology* (4), *Contemporary Educational Psychology* (2), *Journal of Experimental Education* (2), *Written Communication* (1) y *Applied Cognitive Psychology* (1) y el grupo de Fantuzzo en las revistas *Journal of Educational Psychology* (1) y *Journal of Social Behavior and Personality* (1). También Doymuş y sus colegas han difundido sus trabajos en revistas muy variadas como *Journal of Chemical Education* (1), *Research in Science Education*

(1), *Eurasian Journal of Educational Research* (1), *Learning and Instruction* (1) y *Journal of Turkish Science Education* (1).

Por último, un análisis de redes sociales ha permitido identificar los principales grupos y redes de trabajo que han generado una producción científica más activa para cuantificar el grado de *colaboración entre autores* y el grado de *colaboración entre instituciones*.

Con el objetivo de analizar la *colaboración científica entre autores*, se han identificado las coautorías y relaciones de colaboración institucional. De los 90 documentos encontrados, 42 pertenecen a un solo autor –15 son tesis y 27 publicaciones– pero para el análisis de colaboración entre autores se han considerado sólo aquellos documentos en coautoría, es decir, los 48 documentos restantes (53%) que tienen al menos 2 autores (ver tabla 56).

Tabla 56: Relación N° de autores por estudios incluidos

N° autores por documento	N° de documentos
1	42
2	18
3	11
4	11
5	1
6	1
7	6
Total	90

Fuente: elaboración propia

Se han registrado un total de **166** firmas pertenecientes a **109** autores distintos, con un número total de coautorías de **340** siendo 3,5 el número medio de firmas de autores por documento.

En la siguiente tabla, la 57, se recoge para los autores con al menos dos trabajos en colaboración el *patrón de colaboración*. Las dos primeras columnas evidencian el número total de documentos y el número total de firmas en ese grupo de estudios y, en la tercera, el índice firmas/trabajo representa el número medio de firmas/documento para los trabajos de cada autor.

En la columna número de colaboradores se recoge el número de investigadores diferentes que han participado en el conjunto de trabajos, es un indicador del tamaño del grupo de trabajo de cada autor. La última columna, refleja la media de autores por



trabajo al medir el grado de colaboración de cada autor, relacionando el tamaño del grupo de trabajo con la producción.

El autor con mayor grado de colaboración es **Johnson, Roger** con una puntuación de 5 quien tiene entre sus trabajos documentos con un gran número de autores; en contraposición con **Doymuş, Kemal** con un grado de 0,8, valor que indica que su grupo de trabajo es más “cerrado”.

Tabla 57: Resultados patrón de colaboración entre autores

Nº	Autor	Nº documentos	Nº firmas	Índice firmas/trabajo	Nº colaboradores	Nº coautorías	Media* colaboradores/documentos
1	Dansereau, D.	10	55	5,50	13	45	1,40
2	O'Donnell, A.	7	43	6,14	9	36	1,43
3	Larson, C.	6	36	6	9	30	1,67
4	Lambiotte, J.	6	38	6,33	9	32	1,67
5	Rocklin, T.	6	36	6	8	30	1,50
6	<b>Doymuş, K.</b>	5	12	2,40	3	7	<b>0,80</b>
7	Hythecker, V.	5	35	7	9	30	2
8	Young, M.	4	28	7	8	24	2,25
9	Klein, J.	4	12	3	7	8	2
10	Hall, R.	3	16	5,33	9	13	3,33
11	Şimşek, U.	3	10	3,33	3	7	1,33
12	Skaggs, L.	3	15	5	8	12	3
13	Karaçöp, A.	3	10	3,33	3	7	1,33
14	<b>Johnson, R.</b>	2	11	5,50	9	9	<b>5</b>
15	Crooks, S.	2	8	4	5	6	3
16	Dimeff, L.	2	8	4	3	6	2
17	Connelly, S.	2	8	4	3	6	2
18	Fantuzzo, J.	2	8	4	3	6	2
19	Riggio, R.	2	8	4	3	6	2
20	Robinson, D.	2	5	2,50	3	3	2

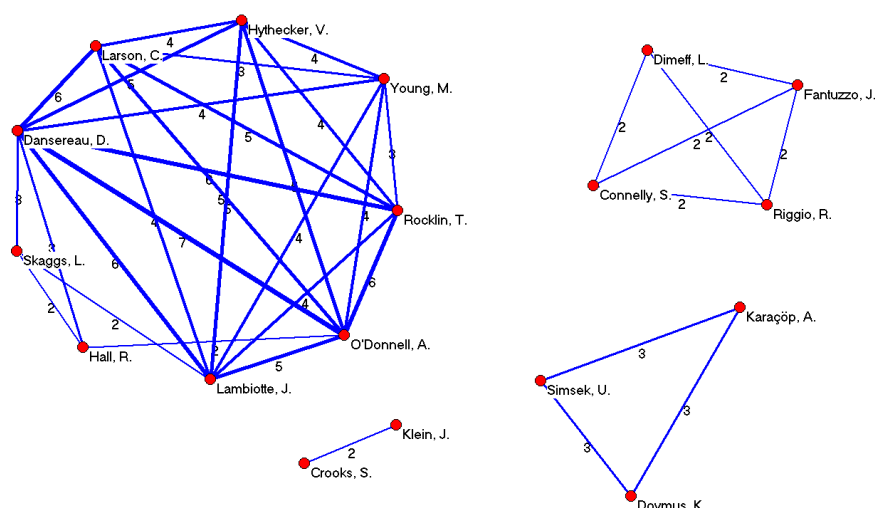
\* Nº de colaboradores + 1 entre el Nº de documentos

Fuente: elaboración propia

La densidad de la red de colaboración entre autores –porcentaje de relaciones existentes entre todas las posibles– es baja, un **5,7%**. Esto indica un gran número de *cluster* de pequeño tamaño formados por un grupo reducido de investigadores. Se trata de parejas de autores con un solo documento en colaboración.

La ilustración 6 representa aquellos *cluster* de autores con **al menos 2 trabajos** en colaboración. Las relaciones de mayor peso aparecen simbolizadas con líneas de mayor grosor. Se han identificado 4 *cluster*, de tamaño máximo 9.

Ilustración 6: Resultados cluster de colaboración entre autores



Nota: Para la representación de la colaboración entre autores se ha utilizado el *software* libre *Pajek* (versión 3.15). Los nodos o vértices de la red se corresponden con los autores y los arcos representan las colaboraciones en forma de coautorías, el peso de los arcos se corresponde con el número de documentos publicados en colaboración.

Fuente: elaboración propia

Continuando con el análisis, la *colaboración entre instituciones* también ha sido estudiada. En la serie de 90 documentos han participado autores de 76 instituciones diferentes, de las cuales 68 estudios pertenecen a una única institución, restando un total de 22 para su análisis (ver tabla 58)

Tabla 58: Resultados N° de instituciones por estudios incluidos

N° instituciones por documento	N° de documentos
1	68
2	19
3	2
5	1
Total	90

Fuente: elaboración propia

Se han identificado un total de **49** firmas pertenecientes a **35** instituciones diferentes, con un número total de coautorías de **66** siendo **2,2** el número medio de instituciones firmantes por documento.

En la siguiente tabla 59 se recoge para las instituciones con al menos dos trabajos en colaboración el *patrón de colaboración* en donde el procedimiento a seguir ha sido el mismo que se ha utilizado para el estudio del grado de colaboración científica entre autores (ver párrafos anteriores).

El grado de colaboración de las instituciones es diferente. Se destaca **Texas Christian University** con el menor grado, un 0,24. Esta universidad tiene un elevado número de publicaciones pero un pequeño grupo de colaboradores, por lo que formaría un grupo de trabajo más “cerrado”. En contraposición, **Password Community Mentoring** en Indianápolis obtiene la media mas alta, con un **2,50**.

Tabla 59: Resultados patrón de colaboración entre instituciones

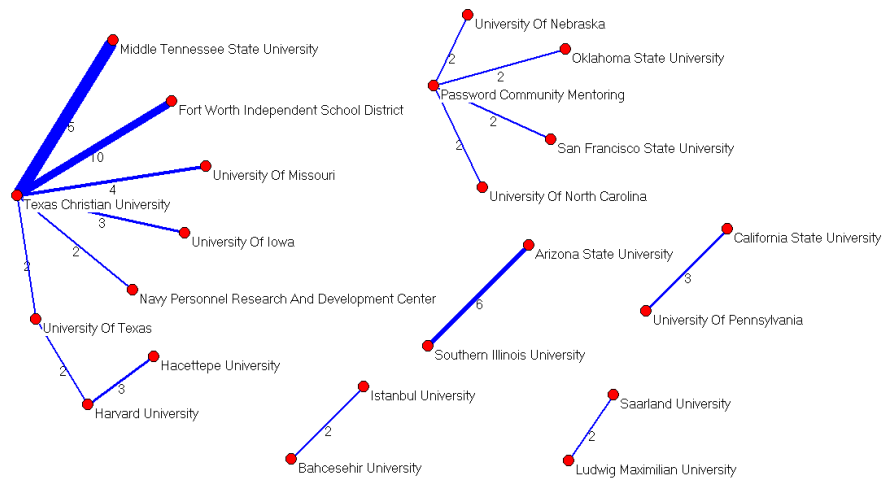
Nº	Instituciones	Nº documentos	Nº firmas	Índice firmas/trabajo	Nº instituciones	Nº coautorías	Media* instituciones/documentos
1	<b>Texas Christian University</b>	29	60	2,07	6	36	<b>0,24</b>
2	Arizona State University	7	14	2,00	2	7	0,43
3	University Of Texas	4	9	2,25	4	6	1,25
4	California State University	3	6	2,00	1	3	0,67
5	Hacettepe University	3	6	2,00	1	3	0,67
6	Middle Tennessee State University	3	6	2,00	1	15	0,67
7	Saarland University	2	4	2,00	1	2	1,00
8	Fort Worth Independent School District	2	4	2,00	1	10	1,00
9	Harvard University	2	4	2,00	2	5	1,50
10	Istanbul University	2	4	2,00	1	2	1,00
11	<b>Password Community Mentoring</b>	2	10	5,00	4	8	<b>2,50</b>
12	Southern Illinois University	2	4	2,00	1	6	1,00

\* N° de instituciones + 1 entre el N° de documentos

Fuente: elaboración propia

La ilustración 7 representa aquellos *cluster* de instituciones **con al menos 2 trabajos en colaboración**. Las relaciones de mayor peso aparecen simbolizadas con líneas de mayor grosor. Se han identificados 6 *cluster*, de tamaño máximo 8.

Ilustración 7: Resultados cluster de colaboración entre instituciones



Nota: Para la representación de la colaboración entre instituciones se ha utilizado el *software* libre *Pajek* (versión 3.15). Los nodos o vértices de la red se corresponden con los autores y los arcos representan las colaboraciones en forma de coautorías, el peso de los arcos se corresponde con el número de documentos publicados en colaboración.

Fuente: elaboración propia

#### 5.1.1.2. Variables metodológicas

La tabla 60 resume las características metodológicas de la muestra, las cuales refieren a aspectos vinculados con la asignación, la situación basal, los sesgos del investigador, los instrumentos de medida y la calidad metodológica de los estudios.

Tabla 60: Resumen características metodológicas de la muestra

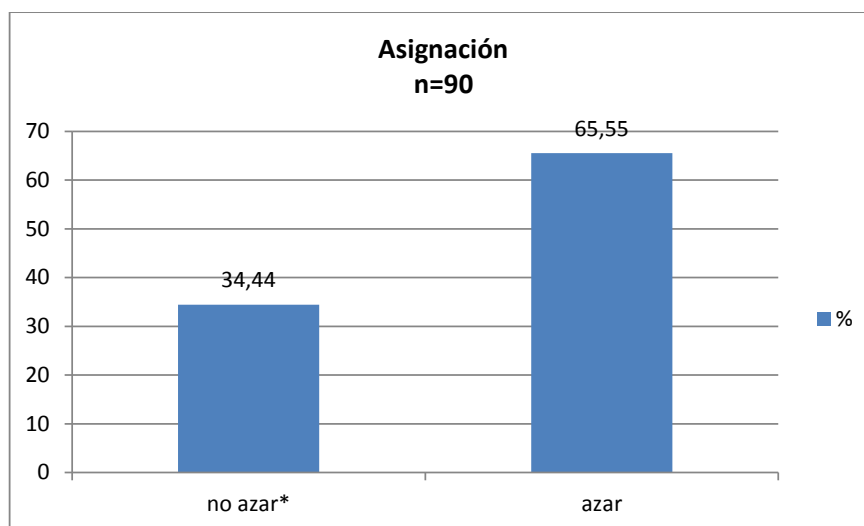
VARIABLES MODERADORAS METODOLÓGICAS		
Asignación	N	%
No azar	31	34,44%
Azar	59	65,55%
Situación basal	N	%
No	18	20%
Si	72	80%
Sesgo del investigador	N	%
Igual investigador	33	36,66%
Diferente investigador	18	20%
Sin especificar	39	43,33%
Instrumentos de medida*	N	%
Estandarizados	47	51,64%
No estandarizados	41	45,05%
Mixtos	3	3,29%
Calidad metodológica	N	%
Baja	10	11,11%
Media	52	57,77%
Alta	28	31,11%

\* suman 91 documentos pero son 90 porque el estudio de McDonald, Larson, Dansereau & Spurlin (1985) que reporta tres *outcomes* ha sido clasificado a la vez en el subgrupo “instrumentos estandarizados” (un *outcome*) e “instrumentos no estandarizados” (dos *outcomes*)

Fuente: elaboración propia

En el gráfico 6 se observa cómo en los estudios primarios, domina la *asignación* al azar de sujetos y/o de grupos al grupo control y experimental (65,55%) y en un menor porcentaje la no asignación aleatoria (34,44%) (ver gráfico 6).

Gráfico 6: Asignación al azar de los sujetos



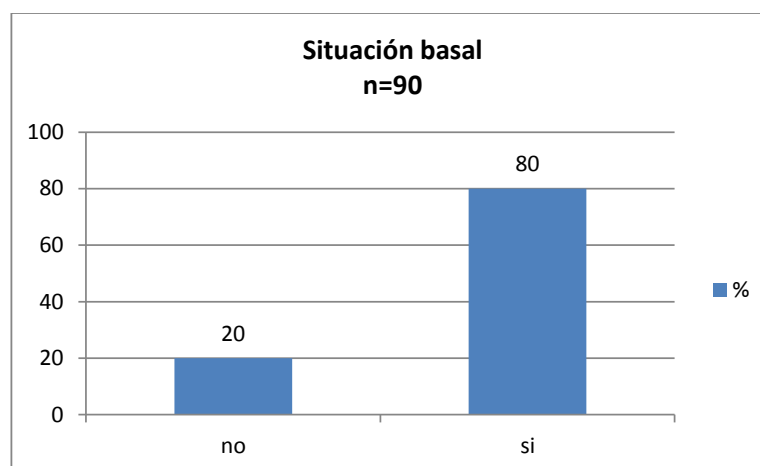
\* en este subgrupo se ha incluido la investigación de Ganter (1994) que es el único estudio de asignación por emparejamiento

Fuente: elaboración propia

Antes de aplicar el programa, en el 80% de los estudios, la *situación basal* se ha considerado, o bien describiendo las características demográficas de los participantes en

cuanto al rendimiento académico en cada grupo, o bien demostrando estadísticamente que el grupo control y experimental son equivalentes. En el 20% restante, el punto de partida de ambos grupos no se ha medido no garantizando así la equivalencia inicial (ver gráfico 7).

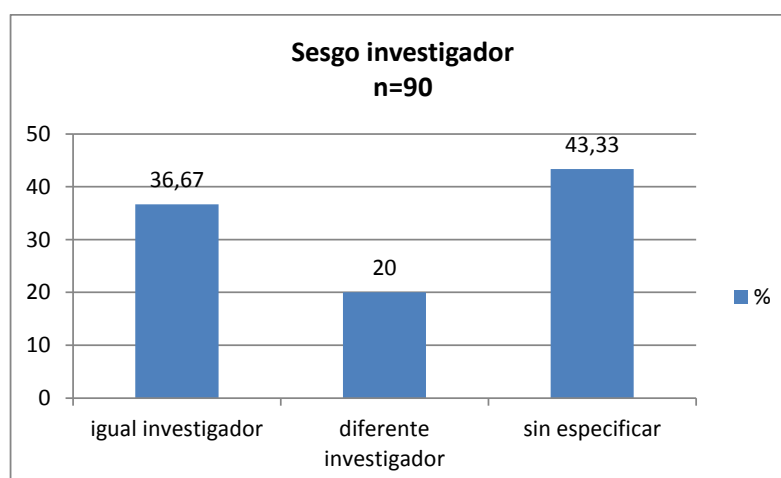
Gráfico 7: Situación basal



Fuente: elaboración propia

Un 43,33% de los estudios primarios no especifica, si son iguales o diferentes investigadores los responsables de la implementación del programa –*sesgo del investigador*– pero cuando se dispone de esta información prevalece que sea el mismo investigador quien se encarga del grupo control y experimental (36,67%) (ver gráfico 8).

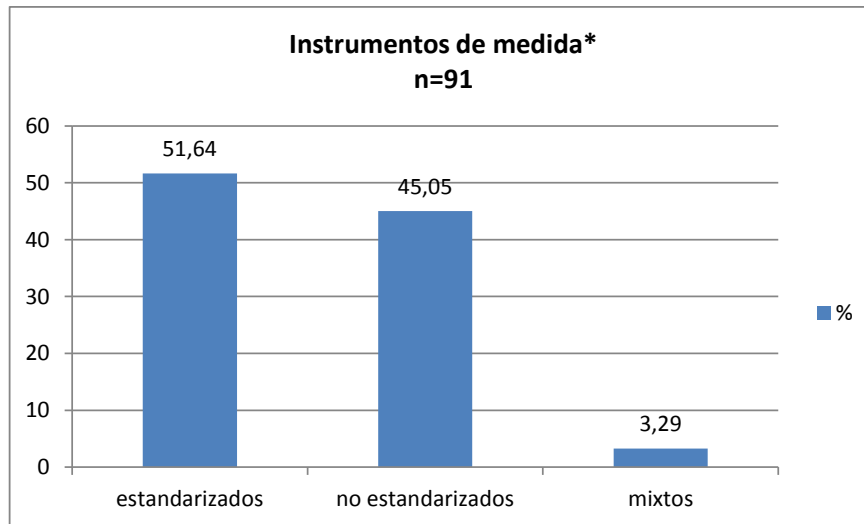
Gráfico 8: Sesgo del investigador



Fuente: elaboración propia

Los instrumentos de medida que se han aplicado en distintas asignaturas para medir el rendimiento de los estudiantes universitarios, son en su mayoría estandarizados (51,64%) aunque también es cierto que un porcentaje elevado, en un 45,05% esta característica no se cumple siendo contados aquellos estudios en donde se emplean a la vez instrumentos estandarizados y no estandarizados (3,29%) (ver gráfico 9).

Gráfico 9: Instrumentos de medida

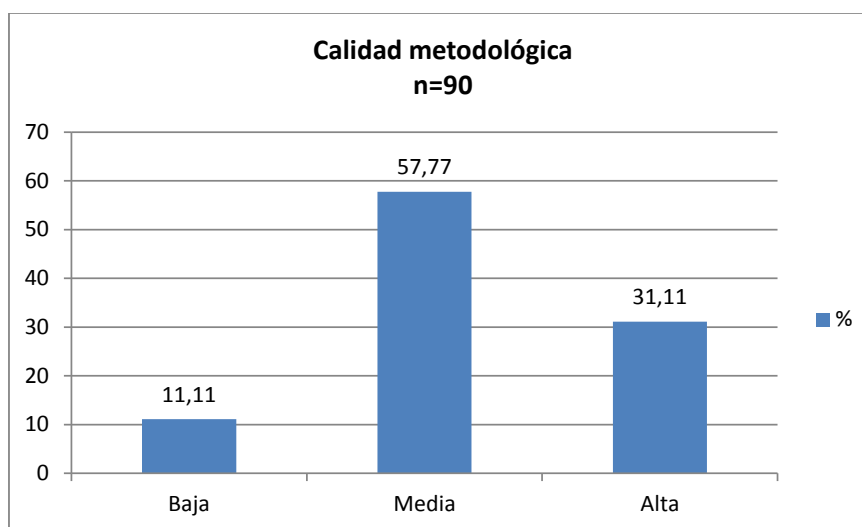


\* suman 91 documentos porque el estudio de McDonald, Larson, Dansereau & Spurlin (1985) que reporta tres *outcomes* ha sido clasificado a la vez en el subgrupo "instrumentos estandarizados" (un *outcome*) e "instrumentos no estandarizados" (dos *outcomes*)

Fuente: elaboración propia

Por último, en cuanto a la *calidad metodológica* de los estudios incluidos sobresale un nivel de calidad medio con un 57,77% del total, un 31,11% de estudios reportan una calidad alta y solo un 11,11% evidencian una baja calidad (ver gráfico 10).

Gráfico 10: Calidad metodológica



Fuente: elaboración propia

### 5.1.1.3. Variables ecológicas

La tabla 61 resume las características sustantivas del objeto de investigación, es decir, aquellas relacionadas con el aprendizaje cooperativo como las asignaturas en donde se ha impartido, los métodos cooperativos empleados, la duración del programa, el número y conformación de los grupos, el tipo de recompensa, la estructura de la tarea y la igualdad de oportunidades para la puntuación.

Tabla 61: Resumen características ecológicas de la muestra

VARIABLES MODERADORAS ECOLÓGICAS		
Áreas de conocimiento	N	%
Agricultura	0	0
Ingeniería, Industria y Construcción	3	3,33%
Salud y Servicios Sociales	3	3,33%
Educación	6	6,67%
No aplica	10	11,11%
Ciencias sociales, Educación Comercial y Derecho	15	16,67%
Humanidades y Artes	16	17,78%
Ciencias	37	41,11%
Métodos cooperativos*	N	%
STL	12	12,63
Cooperative Investigation	4	4,21
Dyadic Methods	14	14,73
Informal Methods	3	3,15
Computer support	17	17,89
Combinados	5	5,26
Sin clasificación	39	41,05
LT	1	1,05
Nº integrantes**	N	%
Dos miembros	29	42,64



Tres miembros	8	11,76
Entre cuatro y seis miembros	31	45,58
<b>Conformación de grupos</b>	<b>N</b>	<b>%</b>
Homogéneo	16	17,78%
Heterogéneo	32	35,56%
Sin especificar	42	46,67%
<b>Duración del programa</b>	<b>N</b>	<b>%</b>
Sesiones	15	16,67%
Entre 1 y 4 semanas	9	10,00%
Entre 5 y 8 semanas	17	18,89%
Entre 9 y 12 semanas	14	15,56%
Entre 13 y 16 semanas	7	7,78%
17 semanas o más	3	3,33%
Sin especificar	25	27,78%
<b>Estructura recompensa*</b>	<b>N</b>	<b>%</b>
Recompensas de grupo basadas en los aprendizajes individuales	12	13,04%
Recompensas de grupo basadas en el producto general	1	1,08%
Sin recompensas de grupo	4	4,34%
No aplica	75	81,52%
<b>Estructura de la tarea*</b>	<b>N</b>	<b>%</b>
Con especialización	8	8,69%
Sin especialización	9	9,78%
No aplica	75	81,52%
<b>Igualdad puntuación*</b>	<b>N</b>	<b>%</b>
Con igualdad	12	13,04%
Sin igualdad	5	5,43%
No aplica	75	81,52%

\* el total sobrepasa los 90 documentos incluidos porque estudios que reportan mas de un *outcome* han sido clasificados a la vez en mas de un subgrupo

\*\* las otras posibilidades de agrupación según el número de integrantes no se han incluido en esta tabla por considerarse muy extensa pero se explicará mas adelante en el texto

Fuente: elaboración propia

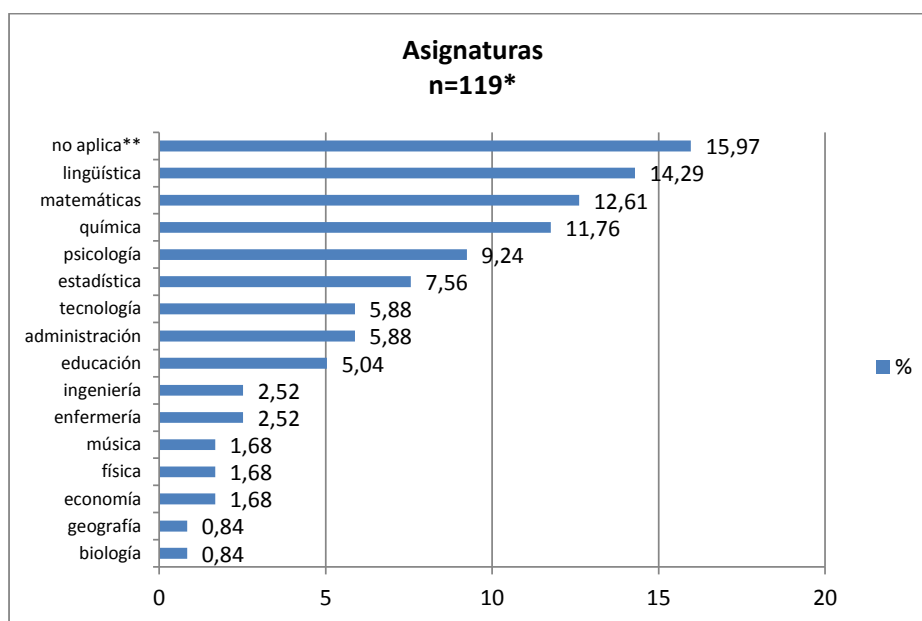
El interés en medir la eficacia del aprendizaje cooperativo en diversas *asignaturas* universitarias, es un reflejo de su versatilidad y así lo demuestra el gráfico 11.

El mayor porcentaje, con un 15,97% lo obtiene la categoría “no aplica”, conformada por estudios que, aunque todos se han realizado con estudiantes de Psicología, el contenido trabajado en las clases no corresponde a esta asignatura. Se tratan de 10 investigaciones –con un total de 19 resultados– en donde se ha implementado el método cooperativo *Scripted Cooperation* para trabajar temas relacionados con Medicina, ciencia ficción, Ecología, Geología y Electrónica, entre otros, razón por la cual se han excluido de esta asignatura y se han categorizado bajo este descriptor.

Dejando a un lado esta categoría, son las otras cinco primeras, las que agrupan más de la mitad de los *outcomes* (55,46%) con asignaturas como Lingüística (14,29%), Matemáticas (12,61%), Química (11,76%), Psicología (9,24%) y Estadística (7,56%). Las siguientes tres con un menor porcentaje corresponden a Tecnología (5,88%), Administración (5,88%) y Educación (5,04%) mientras que aquellas en donde se ha

trabajado con menor frecuencia el aprendizaje cooperativo son Geografía y Biología (0,84%) (ver gráfico 11).

Gráfico 11: Resultados por asignaturas estudios incluidos



\* porcentaje calculado en función del número total de resultados independientes (119 *outcomes*)

\*\* los estudios son Lambiotte, Dansereau, Rocklin, Fletcher, Hythecker, Larson & O'Donnell (1987); Lambiotte, Skaggs & Dansereau (1993); McDonald, Larson, Dansereau & Spurlin (1985); O'Donnell, Dansereau, Hall & Rocklin (1987); O'Donnell, Dansereau, Hythecker, Hall, Skaggs, Lambiotte & Young (1988); O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986); O'Donnell, Dansereau, Rocklin, Hythecker, Lambiotte, Larson & Young (1985); O'Donnell, Dansereau, Rocklin, Larson, Hythecker, Young & Lambiotte (1987); O'Donnell, Larson, Dansereau & Rocklin (1986); y Rewey, Dansereau, Skaggs, Hall & Pitre (1989).

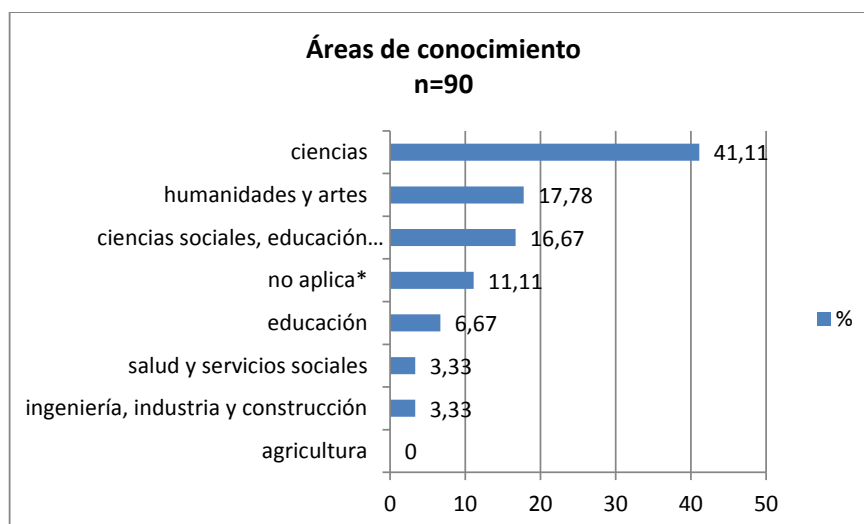
Fuente: elaboración propia

Cuando se agrupan las asignaturas por *áreas de conocimiento* según los criterios establecidos por la Clasificación Internacional Normalizada de la Educación (CINE) de la UNESCO (2011), se observa en el gráfico 12 como el mayor número de documentos se ubica bajo la categoría “Ciencias” con un 41,11% conformada por las Ciencias de la vida, las Ciencias Físicas, Matemáticas, Estadística y Tecnología; seguido de estudios en “Humanidades y Artes” en donde el aprendizaje cooperativo se ha aplicado en asignaturas como Música y Lingüística y, muy de cerca, con un 17,78% “Ciencias Sociales, Educación Comercial y Derecho” con un 16,67% con disciplinas como Administración, Economía y Psicología.

A este listado se suma con un menor porcentaje “Educación” con materias en donde se enseña a docentes en formación Educación Física, Educación en Matemáticas, Educación Musical y Educación para la lectura y escritura con solo un 6,67% quedando

en los últimos tres lugares los programas en “Salud y Servicios Sociales” (3,33%), “Ingeniería, Industria y Construcción” (3,33%) y “Agricultura” sin ningún registro (ver gráfico 12).

Gráfico 12: Resultados por clasificación según Educación y formación estudios incluidos

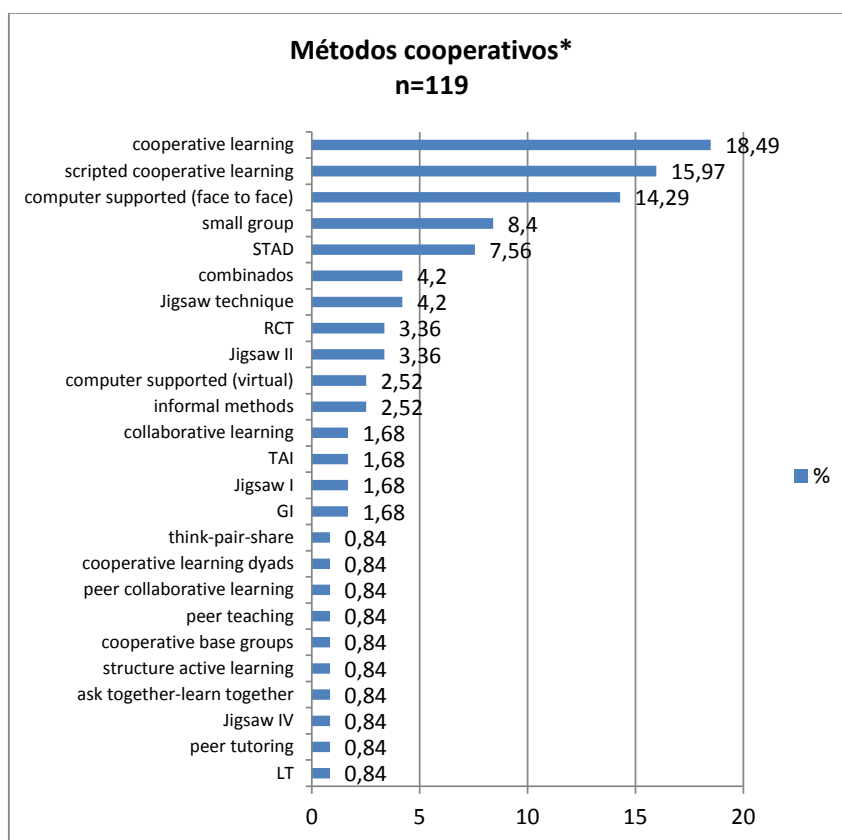


\* incluye los estudios que aunque se han llevado a cabo con estudiantes de Psicología el contenido estudiado no se corresponde con esta área de conocimiento. Estos estudios son Lambiotte, Dansereau, Rocklin, Fletcher, Hythecker, Larson & O'Donnell (1987); Lambiotte, Skaggs & Dansereau (1993); McDonald, Larson, Dansereau & Spurlin (1985); O'Donnell, Dansereau, Hall & Rocklin (1987); O'Donnell, Dansereau, Hythecker, Hall, Skaggs, Lambiotte & Young (1988); O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986); O'Donnell, Dansereau, Rocklin, Hythecker, Lambiotte, Larson & Young (1985); O'Donnell, Dansereau, Rocklin, Larson, Hythecker, Young & Lambiotte (1987); O'Donnell, Larson, Dansereau & Rocklin (1986); y Rewey, Dansereau, Skaggs, Hall & Pitre (1989).

Fuente: elaboración propia

La pluralidad de *métodos cooperativos* es otra muestra de la versatilidad de la que también se hace alusión en la literatura (ver gráfico 13):

Gráfico 13: Resultados métodos cooperativos estudios incluidos



\* porcentaje calculado en función del número total de resultados independientes (119 *outcomes*)

Fuente: elaboración propia

Sin embargo, aunque todos cumplen con los elementos esenciales del aprendizaje cooperativo y aparecen en los estudios primarios bajo el nombre de métodos cooperativos, no todos lo son en sí mismos.

De total de *outcomes*, solo un 61,34% cumple con las exigencias de un método cooperativo (ver tabla 62) siendo el trabajo por parejas *Scripted Cooperation* (15,97%) y el uso del aprendizaje cooperativo con el apoyo de la tecnología *Computer Supported (face to face)* (14,29%) los que predominan, seguidos no tan de cerca por el STAD con un 7,56%.

El resto, un 38,66% se sustenta en los principios del aprendizaje cooperativo pero su técnica no es clara. Tal es el caso de los pseudométodos por parejas que se agrupan en la tabla 62 bajo la denominación *Peer Collaborative Learning*, *Peer Tutoring* y *Cooperative Learning Dyads* o aquellos sin una clasificación específica como *Small Group*, *Collaborative Learning*, *Cooperative Learning* y *Structure Active Learning* o el

tipo de aprendizaje *Cooperative Base Groups* u otras variantes que no son tan conocidas, por ejemplo *Jigsaw Technique*, *Jigsaw IV* y *Ask Together–Learn Together*

Tabla 62: Resultados métodos cooperativos con y sin clasificación estudios incluidos

Con clasificación	N*	%*	Sin clasificación	N*	%*
Scripted Cooperation	19	15,97%	Cooperative Learning	22	18,49%
Computer Support (face to face)	17	14,29%	Small Group	10	8,40%
STAD	9	7,56%	Jigsaw Technique	5	4,20%
Combinados	5	4,20%	Collaborative Learning	2	1,68%
RCT	4	3,36%	Ask Together–Learn Together	1	0,84%
Jigsaw II	4	3,36%	Cooperative Base Groups	1	0,84%
Informal Methods	3	2,52%	Cooperative Learning Dyads	1	0,84%
Computer Support (virtual)	3	2,52%	Peer Collaborative Learning	1	0,84%
GI	2	1,68%	Structure Active Learning	1	0,84%
Jigsaw I	2	1,68%	Peer Tutoring	1	0,84%
TAI	2	1,68%	Jigsaw IV	1	0,84%
LT	1	0,84%			
Peer Teaching	1	0,84%			
Think–Pair–Share	1	0,84%			
Total	73	61,34%	Total	46	38,66%

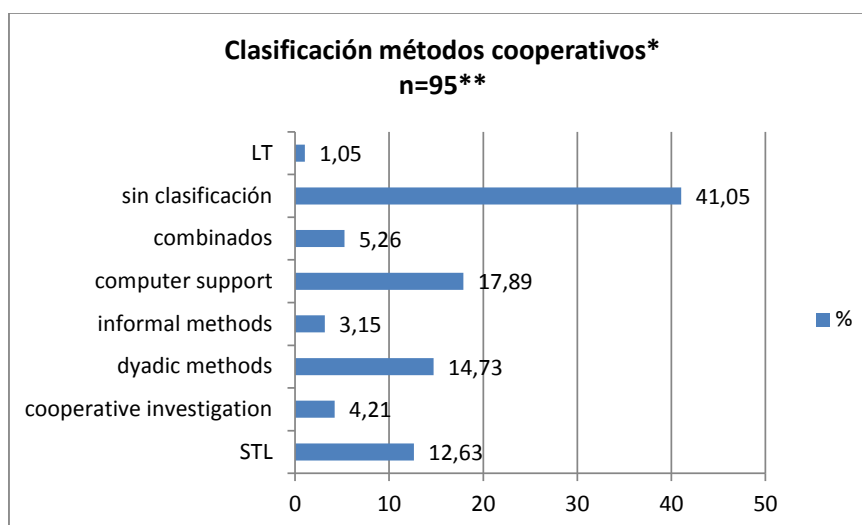
\* calculado a partir de 119 *outcomes*

Fuente: elaboración propia

Por consiguiente, si solo el 61,34% de los 119 *outcomes* cumplen con las exigencias propias de cada técnica, únicamente este porcentaje –lo que equivale al 41,05% de los 90 documentos incluidos– se pueden agrupar a su vez según las *clasificaciones propuestas por los expertos* por tratarse de “verdaderos métodos cooperativos”.

En el gráfico 14 se observa cómo se agrupan estas técnicas cooperativas en donde se evidencia una preferencia en el uso de la tecnología como apoyo a la cooperación *Computer Support* (17,89%), el trabajo por parejas *Dyadic Methods* (14,73%) y los métodos que se aglutinan bajo la categoría STL (12,63%) siendo los menos utilizados *Informal Methods* (3,15%) y LT (1,05%).

Gráfico 14: Clasificación de métodos cooperativos incluidos



\* la clasificación es la siguiente:

STL:STAD + TAI + *Jigsaw II*

*Cooperative Investigation*: GI + *Jigsaw I*

*Dyadic Methods*: *Scripted Cooperation*, RPT, *Peer Teaching + Think-Pair-Share*

*Informal methods*: *Spontaneous Group Discussion*, *Numbered Heads Together*, *Team Product*, *Cooperative Review*

*Computer support*: *face to face + virtual*

Combinados: STL + PBL; *Jigsaw I + Ask Together-Learn Together*; STAD + *Web-Based Collaborative Learning Instruction* (face to face); *Informal Cooperative Learning + Cooperative Learning + Cooperative Base Groups*; *Reciprocal Peer Tutoring + Educational Technology Course* (face to face)

\* suman 95 documentos porque los siguientes estudios al reportar cada uno dos *outcomes* han sido clasificados a la vez en mas de un grupo:

*Computer Support* + sin clasificación: AbuSeileek (2007), Kerr & Murthy (1994) y O'Brien & Peters (1994)

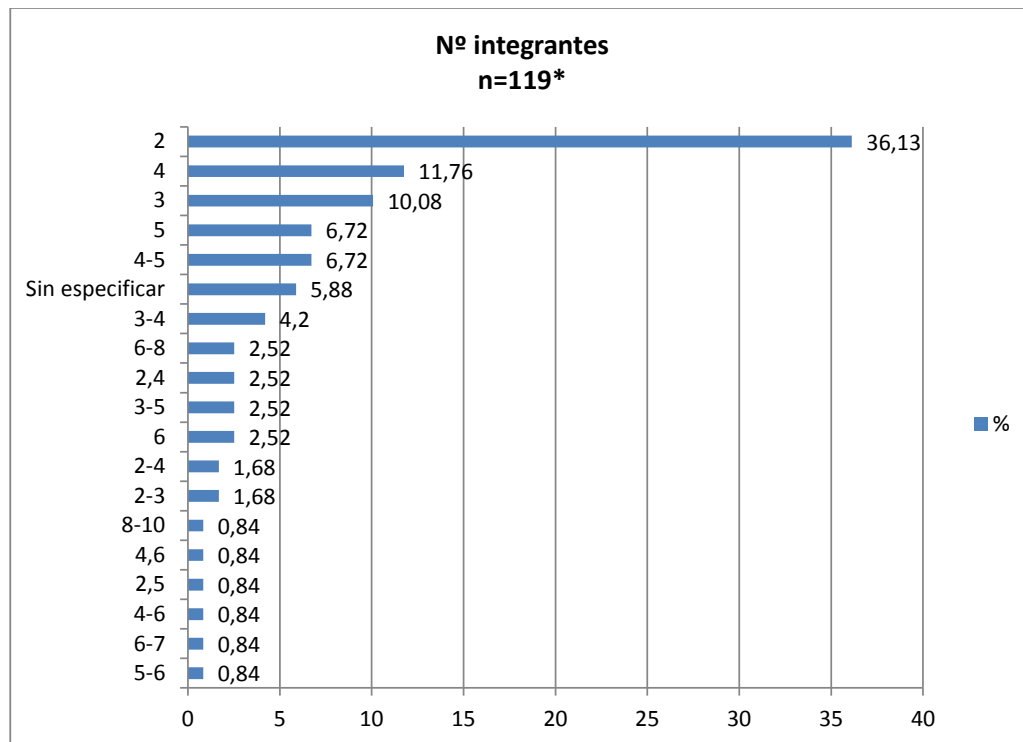
STL + *Cooperative Investigation*: Jalilifar (2010)

*Cooperative Investigation* + sin clasificación: Koç, Doymuş, Karaçöp & Şimşek (2010)

Fuente: elaboración propia

El número de integrantes por grupos cooperativos es variopinto. Así se demuestra en el gráfico 15 en donde la diversidad de combinaciones evidencian distintos criterios para la agrupación de los miembros. Las primeras cinco categorías representan el 71,41% de los *outcomes*, en donde el trabajo por parejas es el que se ha utilizado con mayor frecuencia (36,13%), seguido por el trabajo en grupos de cuatro (11,76%), de tres (10,08%), de cinco (6,72%) y de entre cuatro y cinco (6,72%).

Gráfico 15: Resultados N° de integrantes estudios incluidos

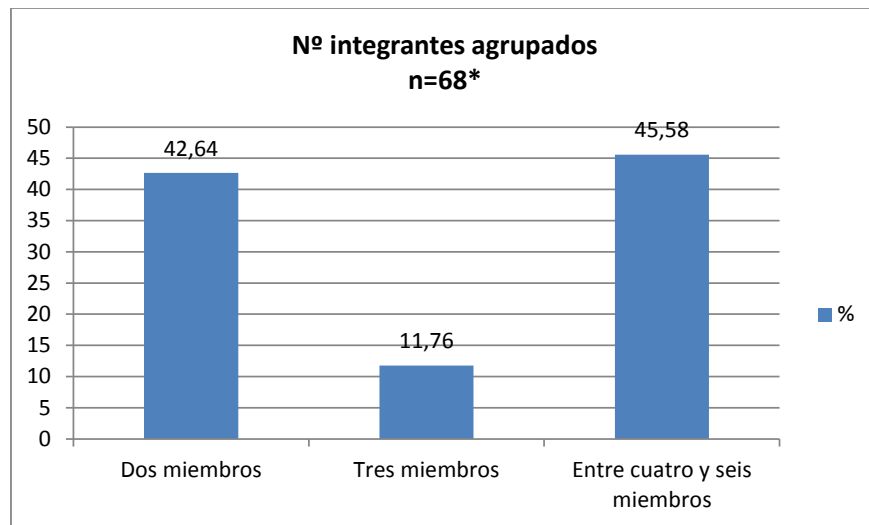


\* porcentaje calculado en función del número total de resultados independientes (119 *outcomes*)

Fuente: elaboración propia

Cuando se analizan solamente las agrupaciones mas frecuentes según el número total de documentos, los grupos conformados entre 4 y 6 estudiantes ocupan el 45,58% pero también el trabajo por parejas prevalece con un 42,64% siendo los tríos los que menos aparecen (11,76%) (ver gráfico 16).

Gráfico 16: Integrantes agrupados

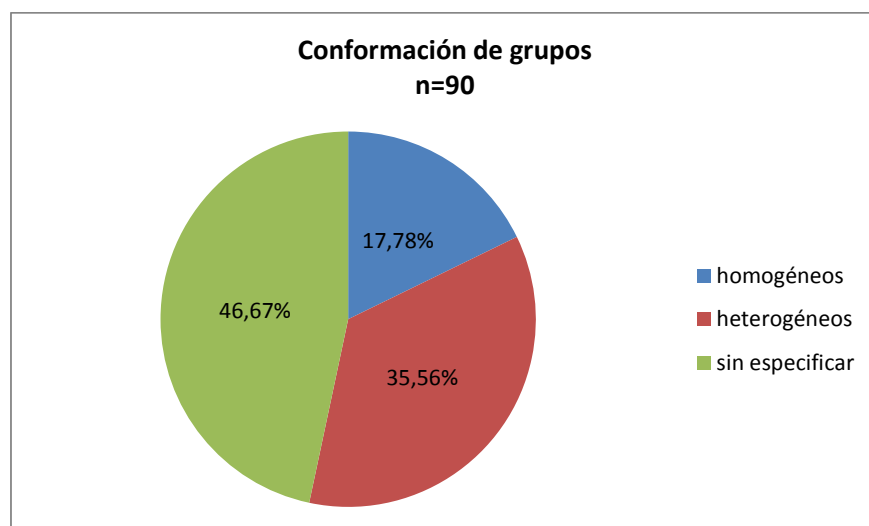


\* suman 68 documentos pero son 67 porque el estudio de O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986) que reporta tres *outcomes* ha sido clasificado a la vez en el subgrupo “dos miembros” (un *outcome*) y en el subgrupo “tres miembros” (2 *outcomes*)

Fuente: elaboración propia

También la *conformación de grupos* es otra característica a considerar para la configuración de las agrupaciones aunque un 46,67% de los estudios primarios no lo especifica y, de aquellos en donde sí aparece esta información, prevalecen los grupos heterogéneos (35,56%) frente a los grupos homogéneos (17,78%) (ver gráfico 17)

Gráfico 17: Conformación de grupos

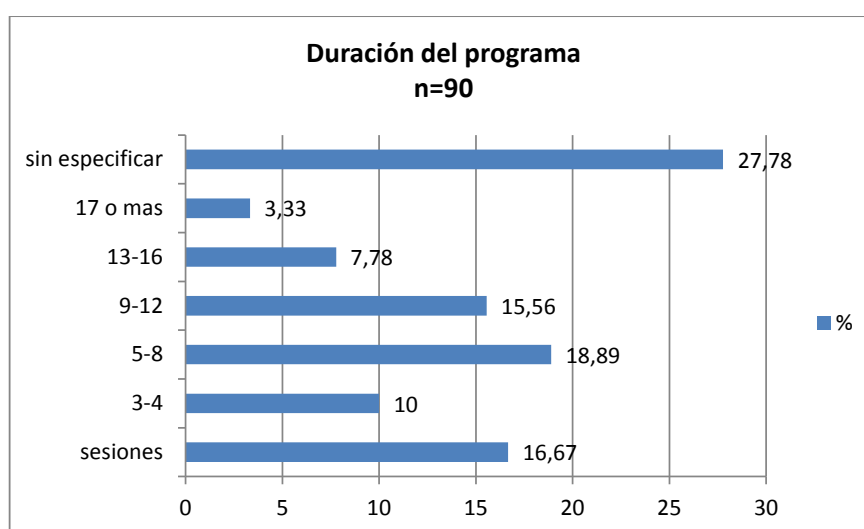


Fuente: elaboración propia



La *duración del programa* abarca un rango de tiempo muy amplio que va desde algunas sesiones hasta casi un semestre de clase. No obstante, en un 27,78% solo se especifica que el programa se ha implementado durante el semestre académico pero no ofrece detalles en cuanto a número, frecuencia y duración. Predominan las intervenciones que abarcan alrededor de la mitad del semestre, entre 5 y 8 semanas (18,89%) y las muy cortas de dos o tres sesiones (16,67%) aunque también resaltan intervenciones mas largas, entre 9 y 12 semanas (15,56%) y escasean aquellas que duran entre 13 y 16 semanas (7,78%) y 17 semanas o mas (3,33%) (ver gráfico 18).

Gráfico 18: Duración del programa



Fuente: elaboración propia

La *estructura de recompensa* –recompensas de grupo basadas en los aprendizajes individuales, recompensas de grupo basadas en el producto grupal o sin recompensas de grupo–, la *estructura de la tarea* –con o sin especialización de la tarea– y la *igualdad de oportunidades para la puntuación grupal* –con o sin igualdades de oportunidades–, son características que diferencian la dinámica interna de los métodos cooperativos STAD, TGT, TAI, CIRC, *Jigsaw I*, *Jigsaw II*, LT y GI (ver gráfico 19).

En general, el número de estudios tan bajo en estas tres características obedece a que solamente pueden incluirse para el análisis los métodos que conforman estas clasificaciones. Por ejemplo no se encontraron estudios con métodos TGT y CIRC razón por la cual en la categoría “no aplica” se agrupan todos los restantes.

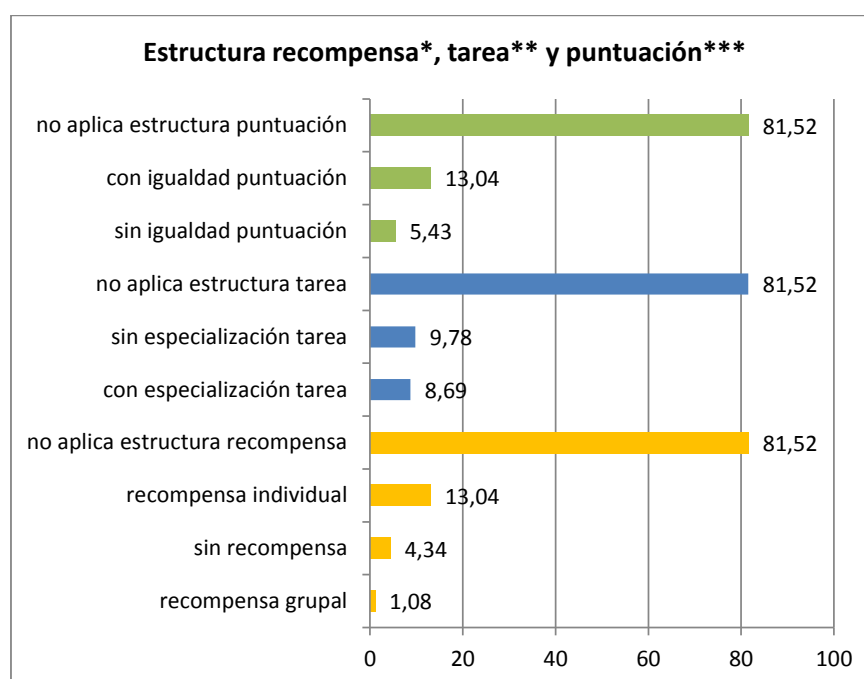
De esta forma, se observa en el gráfico 19 cómo los métodos cooperativos que emplean una *estructura de recompensas*, dominan aquellos en donde se ofrecen retribuciones

individuales (13,04%), luego los que no se sustentan en recompensas (4,34%) y en último lugar, los que utilizan recompensas de grupo basadas en el producto grupal (1,08%).

En cuanto a la *estructura de la tarea*, los porcentajes son muy parecidos entre los métodos con especialización de la tarea (8,69%) de aquellos que se caracterizan por la no existencia de una división del material (9,78%).

Por último, en la *igualdad de oportunidades para la puntuación*, priman los métodos cooperativos que exigen que cada uno de los integrantes del grupo se esfuerce al máximo para lograr la mejor puntuación grupal, con un 13,04% en contraposición con un 5,43% para aquellos que no exigen esta condición (ver gráfico 19).

Gráfico 19: Diferencias entre recompensa, tarea y puntuación



\* los subgrupos que conforman la “estructura de recompensa” –sin contar los que no aplican–suman 31 documentos pero son 30 porque el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado a la vez en mas de un subgrupo: “recompensa individual” (un *outcome*) y “sin recompensa” (dos *outcomes*)

\*\* los subgrupos que conforman la “estructura de la tarea” –sin contar los que no aplican–suman 92 documentos pero son 90 porque los siguientes estudios al reportar cada uno dos *outcomes* han sido clasificados a la vez en mas de un subgrupo:

- alta + baja: Jalilifar (2010)

- alta + no aplica: Koç, Doymuş, Karaçöp & Şimşek (2010)

\*\*\* los subgrupos que conforman la “igualdad de oportunidades para la puntuación” –sin contar los que no aplican– suman 92 documentos pero son 90 porque los siguientes estudios al reportar cada uno dos *outcomes* han sido clasificados a la vez en mas de un subgrupo:

- si + no: Jalilifar (2010)

- no + no aplica: Koç, Doymuş, Karaçöp & Şimşek (2010)

Fuente: elaboración propia

## 5.1.2. Síntesis cuantitativa

### 5.1.2.1. Magnitud del efecto global: el rendimiento

La magnitud del efecto global sobre el rendimiento académico en estudiantes universitarios es estadísticamente significativa y positiva a favor del aprendizaje cooperativo frente al aprendizaje individual. La diferencia media estandarizada de Hedges y Olkin,  $d$ , indica que ese efecto global no solo es significativo sino que su magnitud es similar cuando se comparan los datos globales bajo el modelo de efectos aleatorios con un valor medio de  $d=0,56$  y el modelo de efectos fijos con una  $d=0,57$ .

Siguiendo la clasificación de Cohen, donde  $d=0,2$ ,  $d=0,5$  y  $d=0,8$ , son magnitudes del efecto bajo, medio y alto, respectivamente, se observa que en ambos modelos el efecto global estimado confirma una eficacia de los métodos de aprendizaje cooperativo frente al método de aprendizaje tradicional de tamaño medio (ver tabla 63).

Tabla 63: Magnitud del efecto global: modelo efectos fijos y aleatorios

Modelo	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Pruebas homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
ME (modelo de efectos fijos)	90	119	0,56	0,47	0,65	12,24	$p<0,00001$	430,4	73%
ME (modelo de efectos aleatorios)	90	119	0,57	0,53	0,62	24,97	$p<0,00001$	111,2	0%

Fuente: elaboración propia

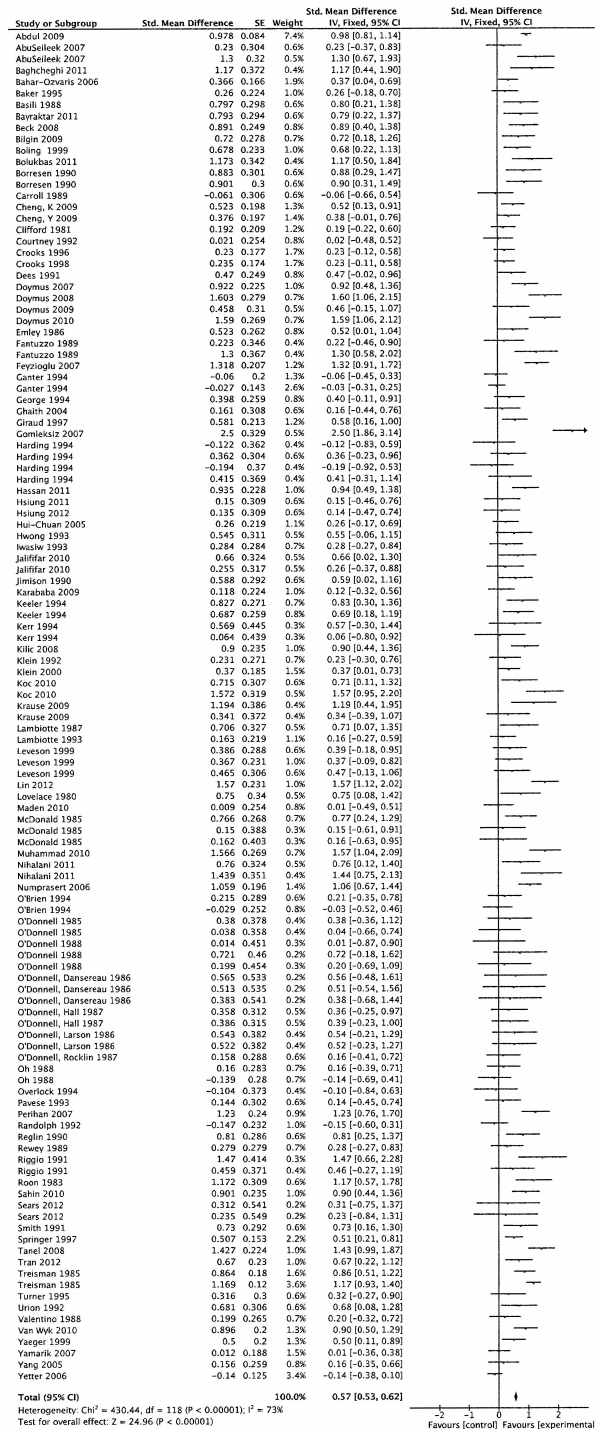
El gráfico *forest plot* representa las medidas del efecto de cada estudio. En orden de izquierda a derecha aparece el primer apellido de cada autor seguido del año de publicación, el tamaño del efecto independiente ( $d$ ), el error estándar (SE), el peso de cada estudio (*weight*), el límite inferior y superior del intervalo de confianza al 95% y la representación gráfica de cada una de las magnitudes del efecto de los 119 *outcomes* provenientes de los 90 documentos primarios incluidos (ver ilustración 8).

Es esta última columna en sí misma en donde se observan las estimaciones del efecto a ambos lados de la línea del efecto nulo que asume el valor de 0 cuando se trabaja con la diferencia de medias. Es el tamaño del área proporcional del cuadrado lo que simboliza la contribución de cada estudio al resultado global. A su vez, cada cuadrado se encuentra en un segmento que representa los extremos de su correspondiente intervalo de confianza. El efecto global también aparece dibujado al final del gráfico en forma de

rombo que igualmente viene acompañado con su correspondiente intervalo de confianza.

Si los límites del intervalo han cruzado la línea del efecto nulo, los estudios no son significativos, pero si estos límites no cruzan la línea, aquellos que han aparecido en el lado derecho de la gráfica son significativos a favor del grupo experimental (aprendizaje cooperativo) y los del lado izquierdo a favor del grupo control (aprendizaje individual).

Ilustración 8: Florest plot: magnitud del efecto de todos los estudios incluidos



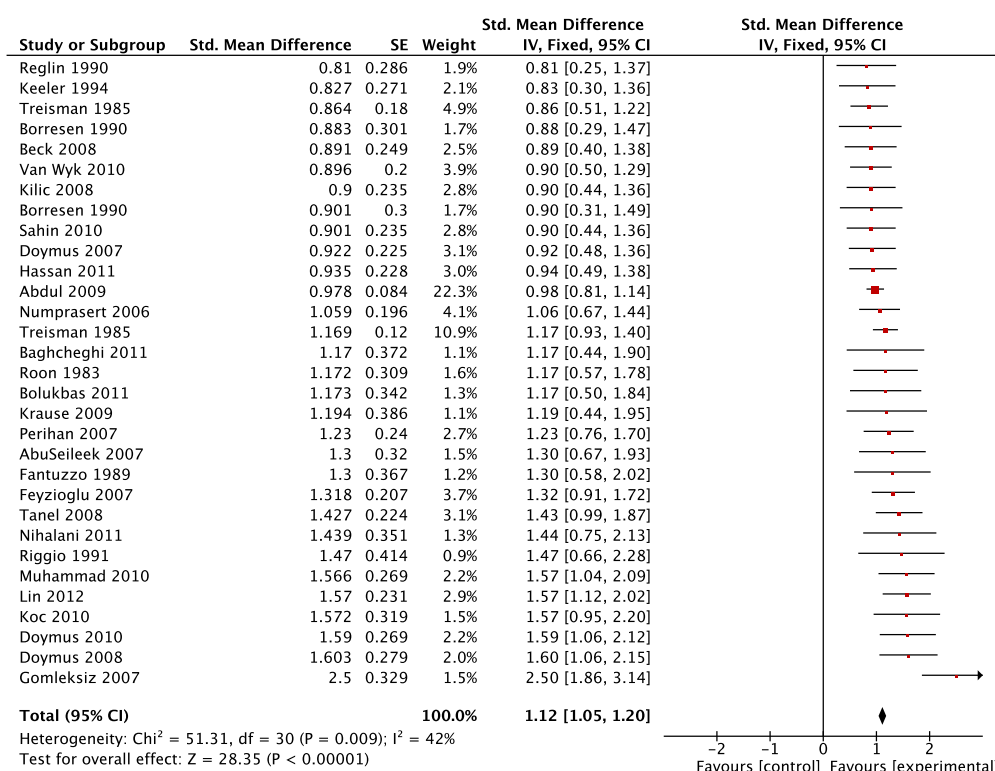
Fuente: elaboración propia

De los 119 resultados (*outcomes*) provenientes de esos 90 documentos estudiados, en 109 de ellos (91,59%) la magnitud del efecto es positiva con efectos individuales que oscilan entre 0,009 y 2,50 y solo en 10 los resultados (8,40%) han sido negativos a favor del aprendizaje individual con efectos que oscilan entre -0,194 y -0,027 (ver ilustración 8).

De estos 109 *outcomes* en donde la magnitud ha sido positiva, el 46,78% ha obtenido un tamaño pequeño (entre 0,009 y 0,47,  $n=51$ ), el 24,77% un tamaño moderado (entre 0,5 y 0,797,  $n=27$ ) y un 28,44% un tamaño alto (entre 0,81 y 2,5,  $n=31$ ) y a su vez, han sido estadísticamente significativos casi la mitad, con 53 resultados (48,62%) a favor del aprendizaje cooperativo con tamaños del efecto que oscilan entre 0,366 hasta 2,5 con efectos grandes (58,49%;  $n=31$ ) seguidos por efectos medianos (35,8%;  $n=19$ ) y en último lugar un porcentaje menor de estudios con un efecto pequeño (5,66%;  $n=3$ ).

En la siguiente ilustración (9), en el gráfico *forest plot* se visualizan esos 31 *outcomes* que han obtenido los tamaños grandes y significativos provenientes de 29 estudios. Si bien se puede identificar que las magnitudes oscilan entre 0,81 y 2,5, es importante tener presente en cuáles de esos estudios el límite inferior del intervalo se acerca o se aleja lo mas posible de la línea del efecto nulo. En el extremo próximo se encuentra Reglin (1990) y en el otro extremo el estudio de Gömleksiz (2007), el cual resalta porque se separa significativamente del efecto nulo.

Ilustración 9: Florest plot: magnitud de efectos grandes y significativos de los estudios incluidos

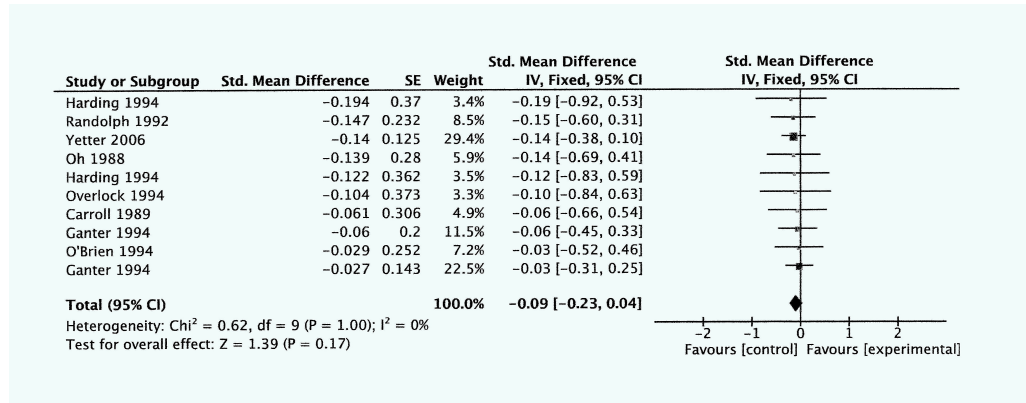


Fuente elaboración propia

Una descripción general de las características de esos 31 *outcomes* indican que la eficacia del aprendizaje cooperativo sobre el individual se ha trabajado con los métodos STL(4), *Dyadic Methods* (2), combinados (1) y *Computer Support* (9) aunque también en 15 no se dispone de información con respecto al método empleado; en las áreas de Ciencias (19), Humanidades y Artes (5), Ciencias Sociales, Educación Comercial y Derecho (3), Salud y Servicios Sociales (2) y Educación (2), provenientes de estudios con niveles de calidad metodológica baja (8), media (20), alta (3) y publicados entre 1983–1999 (9) y entre 1999–2012 (22).

Como se ha mencionado en párrafos anteriores, en cuanto a los resultados que han sido negativos a favor del aprendizaje individual, solo 10 *outcomes* (8,40%) han cumplido con esta condición pero todos no han sido estadísticamente significativos puesto que los intervalos de confianza contienen al 0 (Carrol, 1989; Ganter, 1994; Harding & Fletcher, 1994; O'Brien & Peters, 1994; Oh, 1998; Overlock, 1994; Randolph, 1992; Yetter, Gutkin, Saunders, Galloway, Sobansky & Song, 2006) (ver ilustración 10)

Ilustración 10: Florest plot: magnitud de efectos negativos de los estudios incluidos



Fuente: elaboración propia

#### 5.1.2.2. Análisis de sensibilidad

Como resultado del análisis de sensibilidad –exclusión de estudios uno a uno– se ha calculado la magnitud del efecto global después de eliminar uno de los resultados cada vez.

En la siguiente tabla 64 se muestra la magnitud del efecto para cada estudio, la magnitud del efecto global y el índice de heterogeneidad  $I^2$  al eliminarlo. La última columna indica el cambio relativo en términos porcentuales del tamaño del efecto al eliminarlo respecto del tamaño del efecto global.

Se observa que las magnitudes del efecto al eliminarse se mueven entre 0,55 y 0,57, lo que supone un 2% menos y un 2% más respecto al tamaño del efecto global de 0,56 (modelo de efectos fijos). Tampoco se evidencia un cambio llamativo en el índice de heterogeneidad, que se mueve entre un 70% y un 73%, por lo que se puede decir que ninguno de estos afecta de manera notable al resultado global estimado ( $d=0,56$ ).

En consecuencia, los resultados son bastante robustos, en el sentido de que la estimación de la magnitud del efecto global no cambia de dirección, ni deja de ser estadísticamente significativa



Tabla 64: Análisis de sensibilidad

Análisis de sensibilidad					
Nº	Estudio	Magnitud efecto independiente	Magnitud efecto global al eliminar [IC <sub>95%</sub> ]	I <sup>2</sup> al eliminar	Cambio relativo %
1	Harding, R. & Fletcher, R. (1994)	-0,194	0,57 [0,48; 0,66]	73%	2%
2	Randolph, W. (1992)	-0,147	0,57 [0,48; 0,66]	72%	2%
3	Yetter, G., Gutkin, T., Saunders, A., Galloway, A., Sobansky, R. & Song, S. (2006)	-0,14	0,57 [0,48; 0,66]	71%	2%
4	Oh, H. (1988)	-0,139	0,57 [0,48; 0,66]	72%	2%
5	Harding, R. & Fletcher, R. (1994)	-0,122	0,57 [0,48; 0,66]	73%	2%
6	Overlock, T. (1994)	-0,104	0,57 [0,48; 0,66]	73%	2%
7	Carroll, E. (1989)	-0,061	0,57 [0,48; 0,66]	73%	2%
8	Ganter, S. (1994)	-0,06	0,57 [0,48; 0,66]	72%	2%
9	O'Brien, G. & Peters, J. (1994)	-0,029	0,57 [0,48; 0,66]	72%	2%
10	Ganter, S. (1994)	-0,027	0,57 [0,48; 0,66]	72%	2%
11	Maden, S. (2010)	0,009	0,57 [0,48; 0,66]	73%	2%
12	Yamarik, S. (2007)	0,012	0,57 [0,48; 0,66]	72%	2%
13	O'Donnell, A., Dansereau, D., Hythecker, V., Hall, R., Skaggs, L., Lambiotte, J. & Young, M. (1988)	0,014	0,57 [0,47; 0,66]	73%	2%
14	Courtney, D., Courtney, M. & Nicholson, C. (1992)	0,021	0,57 [0,48; 0,66]	73%	2%
15	O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J., Larson, C. & Young, M. (1985)	0,038	0,57 [0,48; 0,66]	73%	2%
16	Kerr, D. & Murthy, U. (1994)	0,064	0,56 [0,4; 0,66]	73%	0%
17	Karababa, C. (2009)	0,118	0,57 [0,4; 0,66]	73%	2%
18	Hsiung, C. (2012)	0,135	0,57 [0,47; 0,66]	73%	2%
19	Pavese, A. (1993)	0,144	0,57 [0,47; 0,66]	73%	2%
20	McDonald, B., Larson, C., Dansereau, D. & Spurlin, J. (1985)	0,15	0,56 [0,47; 0,66]	73%	0%
21	Hsiung, C. (2011)	0,15	0,57 [0,47; 0,66]	73%	2%
22	Yang, A. (2005)	0,156	0,57 [0,48; 0,66]	73%	2%
23	O'Donnell, A., Dansereau, D., Rocklin, T., Larson, C., Hythecker, V., Young, M. & Lambiotte, J. (1987)	0,158	0,57 [0,47; 0,66]	73%	2%
24	Oh, H. (1988)	0,16	0,57 [0,47; 0,66]	73%	2%
25	Ghaith, G. & El-malak, M. (2004)	0,161	0,57 [0,47; 0,66]	73%	2%
26	McDonald, B., Larson, C., Dansereau, D. & Spurlin, J. (1985)	0,162	0,56 [0,47; 0,65]	73%	0%
27	Lambiotte, J., Skaggs, L. & Dansereau, D. (1993)	0,163	0,57 [0,48; 0,66]	73%	2%
28	Clifford, J. (1981)	0,192	0,57 [0,48; 0,66]	73%	2%
29	Valentino, V. (1988)	0,199	0,57 [0,47; 0,66]	73%	2%
30	O'Donnell, A., Dansereau, D.,	0,199	0,56 [0,47; 0,65]	73%	0%

	Hythecker, V., Hall, R., Skaggs, L., Lambiotte, J. & Young, M. (1988)				
31	O'Brien, G. & Peters, J. (1994)	0,215	0,56 [0,47; 0,66]	73%	0%
32	Fantuzzo, J., Riggio, R., Connelly, S. & Dimeff, L. (1989)	0,223	0,56 [0,47; 0,65]	73%	0%
33	Crooks, S., Klein, J., Jones, E. & Dwyer, H. (1996)	0,23	0,57 [0,47; 0,66]	73%	2%
34	AbuSeileek, A. (2007)	0,23	0,56 [0,47; 0,66]	73%	0%
35	Klein, J. & Pridemore, D. (1992)	0,231	0,56 [0,47; 0,66]	73%	0%
36	Crooks, S., Klein, J., Savenye, W. & Leader, L. (1998)	0,235	0,57 [0,47; 0,66]	73%	2%
37	Sears, D. & Pai, H. (2012)	0,235	0,56 [0,47; 0,65]	73%	0%
38	Jalilifar, A. (2010)	0,255	0,56 [0,47; 0,65]	73%	0%
39	Baker, L. (1995)	0,26	0,56 [0,47; 0,66]	73%	0%
40	Hui-Chuan, L. (2005)	0,26	0,56 [0,47; 0,66]	73%	0%
41	Rewey, K., Dansereau, D., Skaggs, L., Hall, R. & Pitre, U. (1989)	0,279	0,56 [0,47; 0,65]	73%	0%
42	Iwasiw, C. & Goldenberg, D. (1993)	0,284	0,56 [0,47; 0,65]	73%	0%
43	Sears, D. & Pai, H. (2012)	0,312	0,56 [0,47; 0,65]	73%	0%
44	Turner, J. (1995)	0,316	0,56 [0,47; 0,65]	73%	0%
45	Krause, U., Stark, R. & Mandl, H. (2009)	0,341	0,56 [0,47; 0,65]	73%	0%
46	O'Donnell, A., Dansereau, D., Hall, R. & Rocklin, T. (1987)	0,358	0,56 [0,47; 0,65]	73%	0%
47	Harding, R. y Fletcher, R. (1994)	0,362	0,56 [0,47; 0,65]	73%	0%
48	Bahar-Ozvaris, Çuhadaroglu, Turan & Peters (2006)	0,366	0,56 [0,47; 0,66]	73%	0%
49	Leveson, L. (1999)	0,367	0,56 [0,47; 0,65]	73%	0%
50	Klein, J. & Schnackenberg, H. (2000)	0,37	0,56 [0,47; 0,65]	73%	0%
51	Cheng, Y. & Ku, H. (2009)	0,376	0,56 [0,47; 0,65]	73%	0%
52	O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J., Larson, C. & Young, M. (1985)	0,38	0,56 [0,47; 0,65]	73%	0%
53	O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. & Young, M. (1986)	0,383	0,56 [0,47; 0,65]	73%	0%
54	Leveson, L. (1999)	0,386	0,56 [0,47; 0,65]	73%	0%
55	O'Donnell, A., Dansereau, D., Hall, R. & Rocklin, T. (1987)	0,386	0,56 [0,47; 0,65]	73%	0%
56	George, P. (1994)	0,398	0,56 [0,47; 0,65]	73%	0%
57	Harding, R. y Fletcher, R. (1994)	0,415	0,56 [0,47; 0,65]	73%	0%
58	Doymuş, K., Şimşek, U. & Karaçöp, A. (2009)	0,458	0,56 [0,47; 0,65]	73%	0%
59	Riggio, R., Fantuzzo, J., Connelly, S. & Dimeff, L. (1991)	0,459	0,56 [0,47; 0,65]	73%	0%
60	Leveson, L. (1999)	0,465	0,56 [0,47; 0,65]	73%	0%
61	Dees, R. (1991)	0,47	0,56 [0,47; 0,65]	73%	0%
62	Yaeger, P., Marra, R., Costanzo, F. & Gray, G. (1999)	0,5	0,56 [0,47; 0,65]	73%	0%
63	Springer, L. (1997)	0,507	0,56 [0,47; 0,65]	73%	0%
64	O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. &	0,513	0,56 [0,47; 0,65]	73%	0%

	Young, M. (1986)				
65	O'Donnell, A., Larson, C., Dansereau, D. & Rocklin, T. (1986)	0,522	0,56 [0,47; 0,65]	73%	0%
66	Emley, W. (1986)	0,523	0,56 [0,47; 0,65]	73%	0%
67	Cheng, K. (2009)	0,523	0,56 [0,47; 0,65]	73%	0%
68	O'Donnell, A., Larson, C., Dansereau, D. & Rocklin, T. (1986)	0,543	0,56 [0,47; 0,65]	73%	0%
69	Hwong, N., Caswell, A., Johnson, D. & Johnson, R. (1993)	0,545	0,56 [0,47; 0,65]	73%	0%
70	O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. & Young, M. (1986)	0,565	0,56 [0,47; 0,65]	73%	0%
71	Kerr, D. & Murthy, U. (1994)	0,569	0,56 [0,47; 0,65]	73%	0%
72	Giraud, G. (1997)	0,581	0,56 [0,47; 0,65]	73%	0%
73	Jimison, L. (1990)	0,588	0,56 [0,47; 0,65]	73%	0%
74	Jalilifar, A. (2010)	0,66	0,56 [0,47; 0,65]	73%	0%
75	Tran, V. & Lewis, R. (2012)	0,67	0,56 [0,47; 0,65]	73%	0%
76	Boling, N. & Robinson, D. (1999)	0,678	0,56 [0,47; 0,65]	73%	0%
77	Urion, D. & Davidson, N. (1992)	0,681	0,56 [0,47; 0,65]	73%	0%
78	Keeler, C. & Steinhorst, R. (1994)	0,687	0,56 [0,47; 0,65]	73%	0%
79	Lambiotte, J., Dansereau, D., Rocklin, T., Fletcher, B., Hythecker, V., Larson, C. & O'Donnell, A. (1987)	0,706	0,56 [0,47; 0,65]	73%	0%
80	Koç, Y., Doymuş, K., Karaçöp, A. & Şimşek, U. (2010)	0,715	0,56 [0,47; 0,65]	73%	0%
81	Bilgin, I. (2009)	0,72	0,56 [0,47; 0,65]	73%	0%
82	O'Donnell, A., Dansereau, D., Hythecker, V., Hall, R., Skaggs, I., Lambiotte, J. & Young, M. (1988)	0,721	0,56 [0,47; 0,65]	73%	0%
83	Smith, M., Hinckley, C. & Volk, G. (1991)	0,73	0,56 [0,47; 0,65]	73%	0%
84	Lovelace, T. & McKnight, C. (1980)	0,75	0,56 [0,47; 0,65]	73%	0%
85	Nihalani, P., Mayrath, M. & Robinson, D. (2011)	0,76	0,56 [0,47; 0,65]	73%	0%
86	McDonald, B., Larson, C., Dansereau, D. & Spurlin, J. (1985)	0,766	0,56 [0,47; 0,65]	73%	0%
87	Bayraktar, G. (2011)	0,793	0,56 [0,47; 0,65]	73%	0%
88	Basili, P. (1988)	0,797	0,56 [0,47; 0,65]	73%	0%
89	Reglin, G. (1990)	0,81	0,56 [0,47; 0,65]	73%	0%
90	Keeler, C. & Steinhorst, R. (1994)	0,827	0,56 [0,47; 0,65]	73%	0%
91	Treisman, P. (1985)	0,864	0,56 [0,47; 0,65]	73%	0%
92	Borresen, C. (1990)	0,883	0,56 [0,47; 0,65]	73%	0%
93	Beck, L. & Chizhik, A. (2008)	0,891	0,56 [0,47; 0,65]	73%	0%
94	Van Wyk, M. (2010)	0,896	0,56 [0,47; 0,65]	73%	0%
95	Kiliç, D. (2008)	0,9	0,56 [0,47; 0,65]	73%	0%
96	Borresen, C. (1990)	0,901	0,56 [0,47; 0,65]	73%	0%
97	Şahin, A. (2010)	0,901	0,56 [0,47; 0,65]	73%	0%
98	Doymuş, K. (2007)	0,922	0,56 [0,47; 0,65]	73%	0%

99	Hassan, M., Fong, S. & Idrus, R. (2011)	0,935	0,56 [0,47; 0,65]	73%	0%
100	Abdul, S. (2009)	0,978	0,56 [0,47; 0,65]	71%	0%
101	Numprasert, W. (2006)	1,059	0,56 [0,47; 0,65]	72%	0%
102	Treisman, P. (1985)	1,169	0,55 [0,47; 0,64]	71%	-2%
103	Baghcheghi, N., Reza, H. & Rezaei, K. (2011)	1,17	0,56 [0,47; 0,65]	73%	0%
104	Roon, R., Van Pilsum, J., Harris, I., Rosenberg, P., Johnson, R., Liaw, C. & Rosenthal, L. (1983)	1,172	0,56 [0,47; 0,65]	73%	0%
105	Bölükbaş, F., Keskin, F. & Polat, M. (2011)	1,173	0,56 [0,47; 0,65]	73%	0%
106	Krause, U., Stark, R. & Mandl, H. (2009)	1,194	0,56 [0,47; 0,65]	73%	0%
107	Perihan, A. & Kamuran, T. (2007)	1,23	0,56 [0,47; 0,65]	72%	0%
108	Fantuzzo, J., Riggio, R., Connelly, S. & Dimeff, L. (1989)	1,3	0,56 [0,47; 0,65]	73%	0%
109	AbuSeileek, A. (2007)	1,3	0,56 [0,47; 0,65]	72%	0%
110	Feyzioğlu, B., Akçay, H. & Şahin-Pekmez, E. (2007)	1,318	0,55 [0,46; 0,64]	72%	-2%
111	Tanel, Z. & Erol, M. (2008)	1,427	0,55 [0,46; 0,64]	72%	-2%
112	Nihalani, P., Mayrath, M. & Robinson, D. (2011)	1,439	0,56 [0,47; 0,65]	72%	0%
113	Riggio, R., Fantuzzo, J., Connelly, S. & Dimeff, L. (1991)	1,47	0,56 [0,47; 0,65]	73%	0%
114	Muhammad, Z. (2010)	1,566	0,55 [0,46; 0,64]	72%	-2%
115	Lin, Z. (2012)	1,57	0,55 [0,46; 0,64]	72%	-2%
116	Koç, Y., Doymuş, K., Karaçöp, A. & Şimşek, U. (2010)	1,572	0,55 [0,46; 0,64]	72%	-2%
117	Doymuş, K., Karaçöp, A. & Şimşek, U. (2010)	1,59	0,55 [0,46; 0,64]	72%	-2%
118	Doymuş, K. (2008a)	1,603	0,55 [0,46; 0,64]	72%	-2%
119	Gömleksiz, M. (2007)	2,5	0,55 [0,46; 0,63]	70%	-2%

Fuente: elaboración propia

### 5.1.2.3. Heterogeneidad

En el modelo de efectos fijos la prueba de homogeneidad  $Q$  es estadísticamente significativa con  $Q=430,44$ ,  $p<0,00001$  y el porcentaje de variabilidad no explicada por el azar también es alto con un índice de  $I^2=73\%$ , es decir, aproximadamente cerca de un 73% de esa variabilidad no es debida al error de muestreo, rechazándose así la hipótesis nula de homogeneidad mientras que en el modelo de efectos aleatorios se tiene  $Q=111,2$ ,  $p<0,00001$  e  $I^2=0\%$  (ver tabla 63).

Si bien es cierto que el modelo de efectos aleatorios explica que no solo existe variabilidad intraestudios entendida como “las diferencias entre los tamaños del efecto de los estudios individuales se debe solo al hecho de que los estudios utilizan muestras de sujetos diferentes” (Sánchez-Meca, Marín y Huedo, 2006, p.191) sino que además existe una variabilidad interestudios, algunos expertos en la materia consideran que es mucho menos preciso al proporcionar intervalos de confianza mas amplios que el

modelo de efectos fijos. Apuntan los autores que cuando la heterogeneidad es mayor que la que se espera simplemente por azar, puede tener mayor relevancia preguntarse por las causas de esa variabilidad que por el efecto global que se busca estimar (Pértega y Pita, 2005).

Por lo tanto, el análisis ha girado en torno al estudio de las posibles variables moderadoras que expliquen esa heterogeneidad a través de un análisis por subgrupos bajo el modelo de efectos fijos, en el que, las diferencias entre los tamaños del efecto de los estudios individuales se debe a que son muestras de sujetos diferentes (intraestudios) donde “solo será posible generalizar los resultados del meta-análisis a una población de estudios de características idénticas a las de los incluidos” (Catalá y Tobias, 2013, p. 2).

#### 5.1.2.4. Análisis por subgrupos

Debido a la existencia de una alta heterogeneidad en la distribución de las magnitudes del efecto, en el análisis por subgrupos se han estudiado las variables moderadoras que podrían estar incidiendo en los resultados.

Por consiguiente, se ha realizado un análisis univariante mediante el análisis de la varianza ANOVA para descomponer el estadístico  $Q$  en variabilidad entre grupos  $Q_b$  (*between groups*) y variabilidad intragrupos  $Q_w$  (*within groups*) considerando el modelo de efectos fijos, al comparar las características de los estudios previamente codificadas como variables predictoras.

Un resultado estadísticamente significativo en la variabilidad *between groups* ha indicado que existen diferencias entre los grupos, que la variable considerada es fuente de heterogeneidad. Un resultado estadísticamente significativo *within groups* ha señalado que existe variabilidad, es decir, que hay variabilidad dentro de cada subgrupo que no se explica por el azar lo que hace necesario seguir indagando otras variables explicativas.

El análisis de cada una de las variables moderadoras, determinado antes de la recolección de los resultados, se ha estructurado en *variables extrínsecas*, *variables metodológicas* y *variables ecológicas*.

La tabla 65 resume para cada una de estas variables la magnitud del efecto del aprendizaje cooperativo frente al aprendizaje individual en el rendimiento de estudiantes universitarios con su correspondiente intervalo de confianza al 95%. Sin embargo, no todos los subgrupos se han analizado a partir del total de documentos

incluidos (n=90) o bien porque no siempre los estudios han proporcionado información suficiente o bien porque dependiendo de la variable de interés no todos los documentos pueden clasificarse según los criterios establecidos –en especial con subgrupos específicos que pertenecen a la variable ecológica–.

Tabla 65: Resumen magnitudes del efecto por subgrupos según variables extrínsecas, metodológicas y ecológicas

	Variables	Nº de estudios	Nº de resultados independientes	<i>d</i>	IC <sub>95%</sub>		Test of overall effect	
					LI	LS	Z-value	P-value
<b>EXTRÍNECAS</b>	<b>Años de publicación</b>							
	≤ 1999	49	72	0,43	0,37	0,49	13,65	p<0,0001
	≥ 2000	41	47	0,73	0,67	0,8	21,92	p<0,0001
	<b>Lustros</b>							
	80–84	3	3	0,55	0,25	0,85	3,57	0,0004
	85–89	16	28	0,58	0,47	0,69	10,54	p<0,0001
	90–94	21	30	0,3	0,21	0,4	6,06	p<0,0001
	95–99	9	11	0,41	0,28	0,53	6,36	p<0,0001
	00–04	2	2	0,31	0	0,63	1,98	p=0,0473
	05–09	22	24	0,7	0,62	0,78	17,05	p<0,0001
	10–14*	17	21	0,88	0,76	1,01	14,12	p<0,0001
	<b>Décadas</b>							
	80–89	19	31	0,58	0,48	0,68	11,12	p<0,0001
	90–99	30	41	0,34	0,27	0,42	8,7	p<0,0001
	00–09	24	26	0,67	0,6	0,75	17	p<0,0001
	10–19*	17	21	0,88	0,76	1,01	14,12	p<0,0001
	<b>Tipos de documento</b>							
	Reportes técnicos	7	10	0,58	0,44	0,73	7,73	p<0,0001
	Disertaciones/tesis	15	17	0,56	0,44	0,67	9,72	p<0,0001
	Artículos científicos	68	92	0,57	0,52	0,63	21,65	p<0,0001
<b>METODOLÓGICAS</b>	<b>Asignación</b>							
	No azar	30	40	0,72	0,65	0,79	20,39	p<0,00001
	Azar sujetos	32	47	0,4	0,32	0,48	9,82	p<0,0001
	Azar grupos	22	25	0,61	0,51	0,71	11,57	p<0,00001
	Azar sujetos y grupos	5	5	0,74	0,52	0,96	6,5	p<0,0001
	Emparejamiento	1	2	-0,04	-0,27	0,19	-0,33	p=0,743
	<b>Asignación</b>							
	No azar	31	42	0,66	0,59	0,72	19,41	p<0,00001
	Azar	59	77	0,50	0,44	0,56	16,07	p<0,00001
	<b>Situación basal</b>							
	No	18	26	0,51	0,41	0,6	10,64	p<0,0001
	Si	72	93	0,59	0,54	0,64	22,64	p<0,0001
	<b>Sesgo investigador</b>							
	Igual investigador	33	36	0,58	0,50	0,65	15,23	p<0,00001
	Diferente investigador	18	20	0,48	0,38	0,58	9,22	p<0,00001
	<b>Instrumentos de medida</b>							
	Estandarizados	47	57	0,62	0,56	0,69	18,30	p<0,00001
	No estandarizados	41	57	0,55	0,49	0,61	17,26	p<0,00001
	Mixtos	3	5	0,18	-0,09	0,45	1,33	p=0,18

ECOLÓGICAS	<b>Asignaturas</b>							
	Administración	4	7	0,48	0,28	0,68	4,66	p<0,0001
	Lingüística	15	17	0,53	0,4	0,66	8,12	p<0,0001
	Economía	2	2	0,43	0,16	0,7	3,11	p=0,0018
	Educación	6	6	0,7	0,49	0,92	6,47	p<0,0001
	Enfermería	3	3	1,08	0,76	1,4	6,68	p<0,0001
	Estadística	6	9	0,83	0,71	0,96	13,03	p<0,0001
	Física	2	2	1,02	0,64	1,4	5,32	p<0,0001
	Ingeniería	3	3	0,34	0,05	0,63	2,28	p=0,0223
	Matemáticas	10	15	0,58	0,47	0,69	10,03	p<0,0001
	Psicología	9	11	0,25	0,13	0,37	3,95	p<0,0001
	Química	12	14	0,84	0,71	0,97	12,36	p<0,0001
	Tecnología	5	7	0,46	0,27	0,65	4,73	p<0,0001
	No aplica	10	19	0,36	0,20	0,51	4,45	p<0,00001
	<b>Área de conocimiento</b>							
	Ciencias	37	49	0,71	0,64	0,77	21,57	p<0,0001
	Ciencias Sociales, Educación Comercial y Derecho	25	39	0,34	0,25	0,42	7,91	p<0,0001
	Ingeniería, Industria y Construcción	3	3	0,34	0,05	0,63	2,28	p=0,0223
	Humanidades y Artes	16	19	0,52	0,4	0,65	8,12	p<0,0001
	Salud y Servicios Sociales	3	3	1,08	0,76	1,4	6,68	p<0,0001
	Educación	6	6	0,7	0,49	0,92	6,47	p<0,0001
	<b>Métodos cooperativos</b>							
	STL	12	15	0,55	0,41	0,68	7,77	p<0,0001
	Cooperative Investigation	4	4	0,61	0,34	0,89	4,38	p<0,0001
	Dyadic Methods	14	25	0,42	0,28	0,55	6,07	p<0,0001
	Informal Methods	3	3	0,37	0,13	0,62	2,97	p=0,0029
	Computer Support	17	20	0,67	0,56	0,78	11,79	p<0,0001
	Combinados	5	5	0,53	0,35	0,70	5,94	p<0,0001
	<b>Duración del programa</b>							
	Sesiones	15	26	0,3	0,18	0,42	4,91	p<0,00001
	Entre 1 y 4 semanas	9	10	0,79	0,64	0,94	10,14	p<0,00001
	Entre 5 y 8 semanas	17	19	0,72	0,62	0,81	14,32	p<0,00001
	Entre 9 y 12 semanas	14	17	0,5	0,38	0,62	8,34	p<0,00001
	Entre 13 y 16 semanas	7	8	0,51	0,32	0,7	5,36	p<0,00001
	17 semanas o más	3	3	0,55	0,22	0,87	3,28	p=0,001
	<b>Conformación de grupos</b>							
	Homogéneo	16	27	0,41	0,28	0,53	6,45	p<0,00001
	Heterogéneo	32	40	0,56	0,48	0,64	13,18	p<0,00001
	<b>Nº integrantes</b>							
	Dos miembros	29	43	0,45	0,37	0,52	11,81	p<0,00001
	Tres miembros	8	12	0,43	0,26	0,61	4,89	p<0,00001
	Entre cuatro y seis miembros	31	36	0,63	0,55	0,72	14,26	p<0,00001
	<b>Nº integrantes</b>							
	Grupos pares	46	64	0,49	0,43	0,55	15,32	p<0,0001
	Grupos impares	16	20	0,55	0,42	0,67	8,57	p=0,0001
	<b>Estructura recompensa</b>							
	Recompensa de grupos basada en aprendizaje individuales	12	15	0,55	0,41	0,68	7,77	p<0,0001
	Sin recompensa de grupo	4	4	0,61	0,34	0,89	4,38	p<0,0001
	<b>Estructura de tarea</b>							
	Con especialización	8	8	0,88	0,69	1,07	9,09	p<0,00001
	Sin especialización	9	12	0,36	0,21	0,52	4,54	p<0,00001
	<b>Igualdad puntuación</b>							

Con igualdad	12	15	0,55	0,41	0,68	7,77	$p < 0,00001$
Sin igualdad	5	5	0,65	0,40	0,89	5,12	$p < 0,00001$

\* los criterios de inclusión indican estudios hasta el 2012 inclusive

Fuente: elaboración propia

Una mirada general a estos resultados presentados en la tabla anterior (65) evidencia efectos positivos y significativos a favor del aprendizaje cooperativo en todas las variables moderadoras, cuyas magnitudes oscilan entre 0,18 y 1,08, a excepción del subgrupo emparejamiento conformado por un único estudio con un efecto negativo ( $d = -0,04$ ) pero no significativo ( $IC_{95\%} = -0,27; +0,19; p = 0,743$ ) y otros dos subgrupos que aunque reportan tamaños del efecto pequeños tampoco son significativos: el lustro 2000–2004 con una  $d = 0,31$  ( $IC_{95\%} = 0; +0,63; p = 0,0473$ ) y los instrumentos de medida mixtos con una  $d = 0,18$  ( $IC_{95\%} = -0,09; +0,45; p = 0,18$ ).

Cuando se ordenan de mayor a menor los efectos grandes ( $>0,80$ ), un total de siete magnitudes se concentran en variables extrínsecas y ecológicas distribuidas de la siguiente forma: en 17 documentos pertenecientes a los años 2010 y 2012 con una  $d = 0,88$ , en aquellos métodos con estructuras que exigen una especialización de la tarea con un efecto también de 0,88 ( $n = 8$ ) y en asignaturas como Química ( $n = 12, d = 0,84$ ) y Estadística ( $n = 6, d = 0,83$ ) quedando Física ( $d = 1,02$ ), Enfermería ( $d = 1,08$ ) y Salud ( $d = 1,08$ ) con los mayores efectos pero con contados documentos en su haber (2, 3 y 3, respectivamente).

Muy cerca del límite inferior de estos efectos positivos grandes, es decir, próximos al valor de 0,80, se encuentran dos subgrupos interesantes que apuntan también a variables ecológicas a las que se suman variables metodológicas. Por ejemplo, intervenciones cortas entre 1 y 4 semanas obtienen un efecto moderado de 0,79 ( $n = 9$ ) y cuando se asignan al azar a sujetos y a grupos a la situación experimental, el efecto es muy similar con una  $d = 0,74$  ( $n = 5$ ).

En el caso contrario, cuando se ordenan de modo decreciente todas las magnitudes del efecto, los tamaños mas pequeños se aglutinan en instrumentos de medida mixtos con una  $d = 0,18$  –ya mencionado en párrafos anteriores–, en la asignatura de Psicología ( $d = 0,25$ ), la década de los 90–94 ( $d = 0,3$ ) y en sesiones muy breves de dos o tres días ( $d = 0,3$ ).

En síntesis, cuando se comparan los efectos positivos de las variables moderadoras, las ecológicas oscilan entre efectos pequeños y muy grandes (de 0,25 a 1,08), las



metodológicas entre tamaños pequeños y moderados (de 0,4 a 0,66) y las extrínsecas entre magnitudes pequeñas y grandes (de 0,3 a 0,88).

#### 5.1.2.4.1. Variables extrínsecas

Las variables moderadoras extrínsecas refieren a aspectos relacionados con la identificación y publicación de los estudios como son los *años de publicación* y los *tipos de documento*.

##### Años de publicación

Los *años de publicación* en que han sido divulgados los documentos se han agrupado antes del año 1999 y después del año 2000, por lustros y décadas.

Como se observa en la tabla 66, que la magnitud de efecto es mayor ( $d=0,73$ ) en estudios mas recientes, *entre los años 2000 y 2012*, en comparación con estudios mas antiguos, *entre los años 1980 y 1999* ( $d=0,43$ ) pese a que en el primero de los subgrupos la heterogeneidad es mas alta con un 81% y en el segundo, es moderada con un 49%.

Tabla 66: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: 1980-1999 y 2000-2012

Años de publicación 1980-1999 y 2000-2012									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Pruebas homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
$\leq 1999$	49	72	0,43	0,37	0,49	13,65	$p<0,0001$	138,79	49%
$\geq 2000$	41	47	0,73	0,67	0,8	21,92	$p<0,0001$	247,92	81%
Total	90	119							

Fuente: elaboración propia

Con relación a si existe o no diferencia entre grupos ( $Q_b$ ) e intergrupos ( $Q_w$ ) sí se han encontrado diferencias estadísticamente significativas en ambas variabilidades, dato que evidencia heterogeneidad y que a la vez se confirma con una  $I^2$  moderada y alta en cada subgrupo (ver tabla 67).

Recordemos que cuando la prueba de heterogeneidad alcanza significación estadística se interpretará como que los tamaños individuales de los efectos son heterogéneos entre si, lo que se traduce en que la magnitud del efecto medio no los representa bien, por lo que será necesario seguir explorando otras fuentes de variabilidad.

Tabla 67 Resultados del ANOVA variable moderadora: 1980-1999 y 2000-2012

ANOVA		$Q$	P-value
Años de publicación 1980-1999 y 2000-2012	Between groups ( $Q_b$ )	43,74	$p<0,0001$
	Within groups ( $Q_w$ )	386,7	$p<0,0001$
	Total	430,44	$p<0,0001$

Fuente: elaboración propia

Continuando con el análisis, en todos los subgrupos, los resultados por *lustros* de la magnitud del efecto medio son positivos y significativos (ver tabla 68).

Los efectos pequeños se concentran en los estudios publicados en los años 90–94, 95–99 y 00–04 ( $d=0,3$ ,  $0,41$  y  $0,31$ , respectivamente), efectos medianos en los lustros 80–84 y 85–89 ( $d=0,55$ ,  $0,58$ , respectivamente) y efectos mayores en los estudios publicados en los últimos años 05–09 ( $d=0,7$ ) y 10–14 ( $d=0,88$ ) –este último incluye estudios solo hasta el 2012–.

Sin embargo, los lustros 95–99 y 00–04 son los únicos que han obtenido un 0% de heterogeneidad, valor que indica que la variabilidad se debe al azar y no a otras fuentes de explicación pero sin dejar de olvidar que la magnitud del lustro 00–04 no ha sido significativa ( $IC_{95\%}=0$ ;  $+0,63$ ;  $p=0,0473$ ).

Tabla 68: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: lustros

Años de publicación por lustros									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Pruebas homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
80–84	3	3	0,55	0,25	0,85	3,57	$p=0,0004$	7,33	73%
85–89	16	28	0,58	0,47	0,69	10,54	$p<0,0001$	59,58	55%
90–94	21	30	0,3	0,21	0,4	6,06	$p<0,0001$	52,05	44%
95–99	9	11	0,41	0,28	0,53	6,36	$p<0,0001$	5,23	0%
00–04	2	2	0,31	0	0,63	1,98	$p=0,0473$	0,34	0%
05–09	22	24	0,7	0,62	0,78	17,05	$p<0,0001$	171,71	87%
10–14*	17	21	0,88	0,76	1,01	14,12	$p<0,0001$	62,38	68%
Total	90	119							

\* los criterios de inclusión indican estudios hasta el 2012 inclusive

Fuente: elaboración propia

Ambas pruebas de homogeneidad, *between groups*  $Q_b=71,8$ ;  $p<0,0001$  y *within groups*  $Q_w=358,6$ ;  $p<0,0001$  son estadísticamente significativas, lo que quiere decir que hay variabilidad que no se explica por el azar, lo que hace necesario seguir indagando otras variables moderadoras (ver tabla 69).

Tabla 69: Resultados ANOVA variable moderadora: lustros

ANOVA		$Q$	P-value
Años de publicación por lustros	Between groups ( $Q_b$ )	71,8	$p<0,0001$
	Within groups ( $Q_w$ )	358,6	$p<0,0001$
	Total	430,4	$p<0,0001$

Fuente: elaboración propia

Los resultados por *décadas* también evidencian diferencias estadísticamente significativas entre las categorías,  $Q_b=64,8$ ;  $p<0,0001$  (ver tabla 71). Se tienen tamaños del efecto más grandes en la década 00–09 y 10–19 –esta última década incluye estudios solo hasta el 2012– pero también son los que presentan mayor heterogeneidad con un índice de 86% y 68% respectivamente (ver tabla 70).

Tabla 70: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: décadas

Años de publicación por décadas									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
80–89	19	31	0,58	0,48	0,68	11,12	$p<0,0001$	66,95	55%
90–99	30	41	0,34	0,27	0,42	8,7	$p<0,0001$	58,84	32%
00–09	24	26	0,67	0,6	0,75	17	$p<0,0001$	177,51	86%
10–19*	17	21	0,88	0,76	1,01	14,12	$p<0,0001$	62,38	68%
Total	90	119							

\* los criterios de inclusión indican estudios hasta el 2012 inclusive

Fuente: elaboración propia

La prueba de homogeneidad intragrupos también es estadísticamente significativa,  $Q_w=365,7$ ;  $p<0,0001$ , de forma que es necesario seguir estudiando otras fuentes de heterogeneidad (ver tabla 71).

Tabla 71: Resultados ANOVA variable moderadora: décadas

ANOVA		$Q$	P-value
Años de publicación por décadas	Between groups ( $Q_b$ )	64,8	$p<0,0001$
	Within groups ( $Q_w$ )	365,7	$p<0,0001$
	Total	430,4	$p<0,0001$

Fuente: elaboración propia

### Tipos de documento

La variable *tipos de documento*, presenta un comportamiento similar en el estadístico  $Q$  a los años de publicación por décadas pero solo en cuanto a la variabilidad intragrupos por ser significativa  $Q_w=430,3$ ;  $p<0,0001$  no así en la prueba de homogeneidad entre grupos que explica que no existen diferencias significativas  $Q_b=0,105$ ;  $p=0,949$  (ver tabla 72).

Tabla 72: Resultados ANOVA variable moderadora: tipos de documento

ANOVA		$Q$	P-value
Tipos de documento	Between groups ( $Q_b$ )	0,105	p=0,949
	Within groups ( $Q_w$ )	430,3	p<0,0001
	Total	430,4	p<0,0001

Fuente: elaboración propia

Si bien es cierto, que la tabla 73 solo presenta los documentos publicados –no así los no publicados porque no se lograron obtener por falta de respuesta de los investigadores– se evidencia que los tres tipos de documentos reportan magnitudes del efecto similares: los reportes técnicos con una  $d=0,58$ , los artículos en revistas científicas con una  $d=0,57$  y las tesis doctorales con una  $d=0,56$  y a la vez, los dos primeros coinciden con una heterogeneidad moderada y el último alta.

Tabla 73: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: tipos de documento

Tipos de documento									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Reporte técnico	7	10	0,58	0,44	0,73	7,73	p<0,0001	30,77	71%
Disertaciones /tesis	15	17	0,56	0,44	0,67	9,72	p<0,0001	78,57	80%
Artículos de investigación	68	92	0,57	0,52	0,63	21,65	p<0,0001	320,99	72%
Total	90	119							

Fuente: elaboración propia

#### 5.1.2.4.2. Variables metodológicas

Refieren al método y diseño de investigación como son la *asignación*, la *situación basal*, los *sesgos del investigador*, los *instrumentos de medida* y la *calidad metodológica* (este último subgrupo se presenta mas adelante en el apartado 5.1.3).

##### Asignación

Continuando con el análisis de resultados, la *asignación* al azar o no de sujetos y/o grupos es una de las variables moderadoras metodológicas de interés (ver tabla 74).

Llama la atención el hecho de que los 5 documentos en donde la rigurosidad es mayor al asignar al azar tanto a sujetos como a grupos al grupo control y experimental, han obtenido una magnitud del efecto tan alta ( $d=0,74$ ) como las 32 investigaciones en donde no se ha utilizado el azar ( $d=0,72$ ) y que, además en ambos subgrupos la heterogeneidad ha sido moderada–alta (90% y 68%, respectivamente).

El único subgrupo con un documento y dos *outcomes* ha sido la asignación por emparejamiento en donde el efecto ha sido negativo ( $d=-0,04$ ) pero no se han encontrado diferencias significativas entre el grupo control y experimental ( $IC_{95\%} = -0,27; +0,19; p=0,743$ ).

Tabla 74: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: asignación

Asignación									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Pruebas homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
No azar	30	40	0,72	0,65	0,79	20,39	$p<0,00001$	120,13	68%
Azar sujetos	32	47	0,4	0,32	0,48	9,82	$p<0,0001$	103,26	55%
Azar grupos	22	25	0,61	0,51	0,71	11,57	$p<0,00001$	101,57	76%
Azar sujetos y grupos	5	5	0,74	0,52	0,96	6,5	$p<0,0001$	40,17	90%
Emparejamiento	1	2	-0,04	-0,27	0,19	-0,33	0,743	0,02	0%
Total	90	119							

Fuente: elaboración propia

La diferencia del tamaño del efecto en todos los subgrupos ha sido estadísticamente significativa  $Q_b=65,29; p<0,0001$  al igual que la prueba de homogeneidad intragrupos  $Q_w=365,15; p<0,0001$  por lo que puede concluirse que no es suficiente con esta variable para explicar la variabilidad total, siendo necesario dilucidar otras variables moderadoras de la heterogeneidad (ver tabla 75).

Tabla 75: Resultados ANOVA variable moderadora: asignación

ANOVA		$Q$	P-value
Asignación	Between groups ( $Q_b$ )	65,29	$p<0,0001$
	Within groups ( $Q_w$ )	365,15	$p<0,0001$
	Total	430,44	$p<0,0001$

Fuente: elaboración propia

Si se reagrupan los documentos en una categoría mayor (ver tabla 76), conformada por aquellos estudios *sin y con asignación aleatoria* –de sujetos, grupos o ambos– se tiene que en ambos la heterogeneidad es moderada (74% y 71%, respectivamente) y que la magnitud del efecto es menor en aquellos en que se aplica el muestreo probabilístico ( $d=0,50$ ) frente a los que utilizan un muestreo intencional ( $d=0,66$ ).

La prueba de homogeneidad global intragrupos es estadísticamente significativa,  $Q_w=418,61; p<0,0001$  así como la intergrupos  $Q_b=11,83; p<0,0006$  de forma que es necesario estudiar otras fuentes de heterogeneidad (ver tabla 77).

Tabla 76: Resultados de la magnitud del efecto y las pruebas de homogeneidad variable moderadora: asignación azar y no azar

Asignación azar y no azar									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
No azar*	31	42	0,66	0,59	0,72	19,41	p<0,00001	159,30	74%
Azar	59	77	0,50	0,44	0,56	16,07	p<0,00001	259,31	71%
Total	90	119							

\* en este subgrupo se ha incluido la investigación de Ganter (1994) que es el único estudio de asignación por emparejamiento

Fuente: elaboración propia

Tabla 77: Resultados ANOVA variable moderadora: asignación azar y no azar

ANOVA		$Q$	P-value
Asignación azar y no azar	Between groups ( $Q_b$ )	11,83	p=0,0006
	Within groups ( $Q_w$ )	418,61	p<0,0001
	Total	430,44	p<0,0001

Fuente: elaboración propia

### Situación basal

La magnitud del efecto relativa a la *situación basal*, es moderada cuando se demuestra la correspondencia inicial de los grupos ( $d=0,59$ ) pero peculiarmente también cuando no se avala esta igualdad ( $d=0,51$ ). En los dos subgrupos, existen diferencias significativas entre el grupo control y experimental y el porcentaje de heterogeneidad es moderado (72% y 75%, respectivamente) (ver tabla 78).

Tabla 78: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: situación basal

Situación basal									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Pruebas homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
No	18	26	0,51	0,41	0,6	10,64	p<0,0001	100,48	75%
Si	72	93	0,59	0,54	0,64	22,64	p<0,0001	327,39	72%
Total	90	119							

Fuente: elaboración propia

No obstante, con respecto a la variabilidad entre grupos, no se encontraron diferencias estadísticamente significativas al comparar documentos donde se estudia la situación basal de los grupos frente a los que no,  $Q_b=2,6$ ;  $p=0,109$  pero si intragrupos  $Q_w=427,8$ ;  $p<0,0001$  (ver tabla 79).

Tabla 79: Resultados del ANOVA variable moderadora: situación basal

ANOVA		$Q$	P-value
Situación basal	Between groups ( $Q_b$ )	2,6	p=0,109
	Within groups ( $Q_w$ )	427,8	p<0,0001
	Total	430,4	p<0,0001

Fuente: elaboración propia

### Sesgos del investigador

De un total de 90 documentos, solo 51 reportaron información relacionada con el *sesgo del investigador*, donde en 33 estudios el investigador ha sido el mismo tanto en el grupo control como experimental, con un efecto moderado de 0,58 y, en los otros 18 documentos restantes, han sido diferentes investigadores quienes han conducido la situación experimental, con una magnitud baja de 0,48. En los dos subgrupos la heterogeneidad ha sido alta–moderada con un 79% y 67%, respectivamente (ver tabla 80).

Tabla 80: Resultados magnitud del efecto y prueba de homogeneidad variable moderadora: sesgos del investigador

Sesgos del investigador									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Igual investigador	33	36	0,58	0,50	0,65	15,23	$p<0,00001$	164,39	79%
Diferente investigador	18	20	0,48	0,38	0,58	9,22	$p<0,00001$	57,63	67%
Total	51	56							

Fuente: elaboración propia

El análisis *between groups* revela que no se encontraron diferencias estadísticamente significativas entre ambos subgrupos,  $Q_b=2,24$ ;  $p=0,1343$  pero la variabilidad intragrupos es estadísticamente significativa  $Q_w=222,01$ ;  $p<0,0001$  valor que indica que son otras las variables explicativas que podrían estar incidiendo (ver tabla 81).

Tabla 81: Resultados del ANOVA variable moderadora: sesgos del investigador

ANOVA		$Q$	P-value
Sesgos del investigador	Between groups ( $Q_b$ )	2,24	$p=0,1343$
	Within groups ( $Q_w$ )	222,01	$p<0,0001$
	Total	224,26	$p<0,0001$

Fuente: elaboración propia

### Instrumentos de medida

Si los *instrumentos de medida* son estandarizados, la magnitud del efecto es mayor ( $d=0,62$ ) en contraste con instrumentos no estandarizados ( $d=0,55$ ) aunque la heterogeneidad se mantiene en ambos casos moderada (74% y 72% respectivamente).

Si bien es cierto, existen diferencias significativas en todos los subgrupos, cuando se aplican instrumentos tanto validados como no validados, estas diferencias no son significativas +cuando se aplican instrumentos mixtos. En los 3 estudios que reportan

esta situación, la magnitud del efecto ha sido positiva pero baja con una  $d=0,18$  ( $IC_{95\%}=-0,09; +0,45; p=0,18$ ) (ver tabla 82).

Tabla 82: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: instrumentos de medida

Instrumentos de medida									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	$IC_{95\%}$		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Estandarizados	47	57	0,62	0,56	0,69	18,30	$p<0,00001$	218,53	74%
No estandarizados	41	57	0,55	0,49	0,61	17,26	$p<0,00001$	198,02	72%
Mixtos	3	5	0,18	-0,09	0,45	1,33	$p=0,18$	2,67	0%
Total	91*	119							

\* suman 91 documentos pero son 90 porque el estudio de McDonald, Larson, Dansereau & Spurlin (1985) que reporta tres *outcomes* ha sido clasificado a la vez en el subgrupo “instrumentos estandarizados” (un *outcome*) e “instrumentos no estandarizados” (dos *outcomes*)

Fuente: elaboración propia

Sin embargo, existen diferencias significativas entre grupos  $Q_b=11,23$ ;  $p=0,0036$  y también intragrupos  $Q_w=419,21$ ;  $p<0,0001$  lo que significa que al igual que otras variables ya analizadas anteriormente es necesario seguir indagando otras fuentes de variabilidad (ver tabla 83)

Tabla 83: Resultados ANOVA variable moderadora: instrumentos de medida

ANOVA		$Q$	P-value
Instrumentos de medida	Between groups ( $Q_b$ )	11,23	$p=0,0036$
	Within groups ( $Q_w$ )	419,21	$p<0,0001$
	Total	430,44	$p<0,0001$

Fuente: elaboración propia

#### 5.1.2.4.3. Variables ecológicas

Las variables ecológicas o sustantivas refieren al objeto de investigación, en este caso a cómo se han agrupado las asignaturas por *áreas de conocimiento*, los *métodos cooperativos*, la *duración del programa*, la *conformación de los grupos*, el *número de integrantes*, la *estructura de la recompensa*, la *estructura de la tarea* y la *igualdad de oportunidades en la puntuación*.

#### Áreas de conocimiento

En un primer análisis, para el cálculo de la magnitud del efecto según las *asignaturas* en donde se ha impartido el aprendizaje cooperativo, de los 90 estudios, se han excluido tres debido a que cuenta cada uno con un solo *outcome*, los cuales corresponden a



Música (Sears & Pai, 2012), Biología (Randolph, 1992) y Geografía (Numprasert, 2006).

Todos los subgrupos, sin excepción alguna, evidencian la eficacia de la cooperación en las asignaturas estudiadas siendo el efecto estadísticamente significativo entre el grupo control y experimental en cada una de ellas (ver tabla 84).

En Enfermería, Física, Química y Estadística –ordenadas de mayor a menor– la magnitudes del efecto son altas con puntuaciones que van en un rango entre 1,08 y 0,83 pero con un número muy bajo de estudios a excepción de Química que cuenta con 12 documentos ( $d=0,84$ ), Enfermería con tres ( $d=1,08$ ), Física con dos ( $d=1,02$ ) y Estadística con seis estudios ( $d=0,83$ ).

Con un mayor número de estudios, efectos de tamaño medio se presentan en Educación ( $d=0,7$ ), Matemáticas ( $d=0,58$ ) y Lingüística ( $d=0,53$ ) seguidos de magnitudes bajas en Administración ( $d=0,48$ ), Tecnología ( $d=0,46$ ) y Economía ( $d=0,43$ ) ocupando los dos últimos lugares Ingeniería con  $d=0,34$  ( $n=3$ ) y Psicología con  $d=0,25$  ( $n=9$ ).

El porcentaje de heterogeneidad es alto en orden decreciente en Física (92%), Economía, Enfermería (90%), Matemáticas (83%), Lingüística (77%) y Química (74%) e Ingeniería y Administración con un 0%, dato último que indica que la variabilidad se debe al azar y no a otras fuentes de explicación.

Tabla 84: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: asignaturas

Asignaturas									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Administración	4	7	0,48	0,28	0,68	4,66	p<0,0001	2,01	0%
Lingüística	15	17	0,53	0,4	0,66	8,12	p<0,0001	70,03	77%
Economía	2	2	0,43	0,16	0,7	3,11	p=0,0018	10,37	90%
Educación	6	6	0,7	0,49	0,92	6,47	p<0,0001	13,5	63%
Enfermería	3	3	1,08	0,76	1,4	6,68	p<0,0001	12,41	84%
Estadística	6	9	0,83	0,71	0,96	13,03	p<0,0001	17,62	55%
Física	2	2	1,02	0,64	1,4	5,32	p<0,0001	12,38	92%
Ingeniería	3	3	0,34	0,05	0,63	2,28	p=0,0223	1,46	0%
Matemáticas	10	15	0,58	0,47	0,69	10,03	p<0,0001	80,44	83%
Psicología	9	11	0,25	0,13	0,37	3,95	p<0,0001	28,19	65%
Química	12	14	0,84	0,71	0,97	12,36	p<0,0001	50,48	74%
Tecnología	5	7	0,46	0,27	0,65	4,73	p<0,0001	18,31	67%
No aplica*	10	19	0,36	0,20	0,51	4,45	p<0,00001	8,12	0%
Total	87	115							

\* conformada por estudios que, aunque todos se han llevado a cabo con estudiantes de Psicología, el contenido trabajado en las clases no corresponde a esta asignatura. En todos se ha implementado el método *Scripted Cooperation* con temas relacionados con Medicina, ciencia ficción, Ecología, Geología y Electrónica, entre otros. Los estudios son Lambiotte, Dansereau, Rocklin, Fletcher, Hythecker, Larson & O'Donnell (1987); Lambiotte, Skaggs & Dansereau (1993); McDonald, Larson, Dansereau & Spurlin (1985); O'Donnell, Dansereau, Hall & Rocklin (1987); O'Donnell, Dansereau, Hythecker, Hall, Skaggs, Lambiotte & Young (1988); O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986); O'Donnell, Dansereau, Rocklin, Hythecker, Lambiotte, Larson & Young (1985); O'Donnell, Dansereau, Rocklin, Larson, Hythecker, Young & Lambiotte (1987); O'Donnell, Larson, Dansereau & Rocklin (1986); y Rewey, Dansereau, Skaggs, Hall & Pitre (1989).

Fuente: elaboración propia

No obstante, ambas pruebas de homogeneidad *between groups* y *within groups* son significativas lo que significa que las diferencias encontradas no se deben al azar sino a otras fuentes de variabilidad (ver tabla 85)

Tabla 85: Resultados del ANOVA variable moderada: asignaturas

ANOVA		$Q$	P-value
Asignaturas	Between groups ( $Q_b$ )	88,73	p<0,0001
	Within groups ( $Q_w$ )	325,32	p<0,0001
	Total	414,05	p<0,0001

Fuente: elaboración propia

En un segundo análisis (ver tabla 86), con la intención de agrupar las asignaturas por *áreas de conocimiento* –no habiéndose registrado ningún estudio en el área de Agricultura–, de los 90 documentos, la magnitud del efecto mayor se reporta en 3 estudios en el área de Salud y Servicios Sociales –Enfermería– con una  $d=1,08$  pero con un índice de heterogeneidad del 84%, seguido de otros tamaños del efecto igualmente altos como en Ciencias –Biología, Estadística, Física, Geografía, Matemáticas, Química e Informática– con una  $d=0,71$  y un 78% de heterogeneidad y en 6 estudios en Educación –Educación Física, Educación en Matemáticas, Educación

Musical y Educación para la Lectura y Escritura– con una  $d=0,7$  pero con un porcentaje moderado de heterogeneidad, con un 63% (ver tabla 86).

Con 16 estudios, las Humanidades y Artes –Lingüística y Música– han obtenido un efecto medio con  $d=0,52$  ( $I^2=74\%$ ) quedando las Ciencias Sociales, Educación Comercial y Derecho –Administración, Economía y Psicología– y la Ingeniería, Industria y Construcción con las magnitudes del efecto mas bajas (0,33 y 0,34, respectivamente) (ver tabla 86).

Tabla 86: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: clasificación de la Educación y formación

Áreas de conocimiento									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	Q	I <sup>2</sup>
Ciencias	37	49	0,71	0,64	0,77	21,57	$p<0,0001$	217,36	78%
Ciencias Sociales, Educación Comercial y Derecho	15	20	0,33	0,23	0,42	6,55	$p<0,00001$	44,75	58%
Ingeniería, Industria y Construcción	3	3	0,34	0,05	0,63	2,28	$p=0,0223$	1,46	0%
Humanidades y Artes	16	19	0,52	0,4	0,65	8,12	$p<0,0001$	70,46	74%
Salud y Servicios Sociales	3	3	1,08	0,76	1,4	6,68	$p<0,0001$	12,41	84%
Educación	6	6	0,7	0,49	0,92	6,47	$p<0,0001$	13,5	63%
No aplica*	10	19	0,36	0,20	0,51	4,45	$p<0,00001$	8,12	0%
Total	90	119							

\* conformada por estudios que, aunque todos se han llevado a cabo con estudiantes de Psicología, el contenido trabajado en las clases no corresponde a esta asignatura. En todos se ha implementado el método *Scripted Cooperation* con temas relacionados con Medicina, ciencia ficción, Ecología, Geología y Electrónica, entre otros. Los estudios son Lambiotte, Dansereau, Rocklin, Fletcher, Hythecker, Larson & O'Donnell (1987); Lambiotte, Skaggs & Dansereau (1993); McDonald, Larson, Dansereau & Spurlin (1985); O'Donnell, Dansereau, Hall & Rocklin (1987); O'Donnell, Dansereau, Hythecker, Hall, Skaggs, Lambiotte & Young (1988); O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986); O'Donnell, Dansereau, Rocklin, Hythecker, Lambiotte, Larson & Young (1985); O'Donnell, Dansereau, Rocklin, Larson, Hythecker, Young & Lambiotte (1987); O'Donnell, Larson, Dansereau & Rocklin (1986); y Rewey, Dansereau, Skaggs, Hall & Pitre (1989).

Fuente: elaboración propia

Las pruebas de variabilidad entre grupos  $Q_b=62,3$ ;  $p<0,0001$  y variabilidad intragrupos  $Q_w=368,2$ ;  $p<0,0001$  han sido estadísticamente significativas, valores que indican que existen diferencias entre los subgrupos y que la heterogeneidad intraestudios invita a seguir indagando en otras variables explicativas de los tamaños del efecto (ver tabla 87).

Tabla 87: Resultados del ANOVA variable moderadora: áreas de conocimiento

ANOVA		$Q$	P-value
Áreas de conocimiento	Between groups ( $Q_b$ )	62,3	p<0,0001
	Within groups ( $Q_w$ )	368,2	p<0,0001
	Total	430,4	p<0,0001

Fuente: elaboración propia

### Métodos cooperativos

Solo 56 de 90 documentos se han logrado categorizar según la clasificación de *métodos cooperativos* propuesta por los expertos pero de estos, se ha excluido uno para el cálculo de la magnitud del efecto debido a que solo cuenta con un resultado (*Learning Together*: Bayraktar, 2011). Por tanto, el total de documentos es 54 aunque la sumatoria indique 55 porque el estudio de Jalilifar (2010) que tiene dos *outcomes* aparece repetido en los subgrupos STL y *Cooperative Investigation* (ver tabla 88).

En la tabla 88, se evidencia cómo los métodos en parejas *Dyadic Methods* y los informales *Informal Methods* son los subgrupos con menores magnitudes del efecto ( $d=0,42$  y  $d=0,37$ , respectivamente), con los porcentajes mas bajos de heterogeneidad (0% y 45%, respectivamente) y a la vez *Informal Methods* es el subgrupo con el menor número de documentos, solo 3.

El resto de métodos con un efecto moderado se aglutinan en orden decreciente bajo la clasificación *Computer Support* ( $d=0,67$ ), *Cooperative Investigation* ( $d=0,61$ ), STL ( $d=0,55$ ) y cuando se *combinan* varios métodos cooperativos ( $d=0,53$ ).

Es de comentar que los cuatro estudios del subgrupo *Cooperative Investigation* han obtenido un 0% de heterogeneidad mientras que los 12 estudios en donde se ha aplicado el STL y los 5 estudios en donde se han *combinado* varias estrategias cooperativas son los que han obtenido la heterogeneidad mas alta con un 81% respectivamente.

Tabla 88: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: clasificación métodos

Clasificación métodos cooperativos*									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	Q	I <sup>2</sup>
STL	12	15	0,55	0,41	0,68	7,77	p<0,0001	72,45	81%
Cooperative Investigation	4	4	0,61	0,34	0,89	4,38	p<0,0001	1,61	0%
Dyadic Methods	14	25	0,42	0,28	0,55	6,07	p<0,0001	21,50	0%
Informal Methods	3	3	0,37	0,13	0,62	2,97	p=0,0029	3,63	45%
Computer Support	17	20	0,67	0,56	0,78	11,79	p<0,0001	83,08	77%
Combinados	5	5	0,53	0,35	0,70	5,94	p<0,0001	21,01	81%
Total	55**	72							

\* la clasificación es la siguiente:

STL:STAD + TAI + Jigsaw II

Cooperative Investigation: GI + Jigsaw I

Dyadic Methods: Scripted Cooperation, RPT, Peer Teaching + Think-Pair-Share

Informal methods: Spontaneous Group Discussion, Numbered Heads Together, Team Product, Cooperative Review

Computer support: face to face + virtual

Combinados: STL + PBL; Jigsaw I + Ask Together-Learn Together; STAD + Web-Based Collaborative Learning Instruction (face to face); Informal Cooperative Learning + Cooperative Learning + Cooperative Base Groups; Reciprocal Peer Tutoring + Educational Technology Course (face to face)

\*\* suman 55 documentos pero son 54 porque el estudio de Jalilifar (2010) que tiene dos *outcomes* aparece repetido en los subgrupos STL y Cooperative Investigation

Fuente: elaboración propia

Sin embargo, no se encontraron diferencias significativas entre las distintas clasificaciones de los métodos cooperativos  $Q_b=10,3$ ;  $p=0,0661$  pero si intragrupos  $Q_w=203,3$ ;  $p<0,0001$ , de forma que es necesario estudiar otras fuentes de heterogeneidad (ver tabla 89).

Tabla 89: Resultados del ANOVA variable moderadora: clasificación métodos cooperativos

ANOVA		Q	P-value
Clasificación métodos cooperativos	Between groups ( $Q_b$ )	10,3	p=0,0661
	Within groups ( $Q_w$ )	203,3	p<0,0001
	Total	213,6	p<0,0001

Fuente: elaboración propia

### Duración del programa

Solo 65 documentos presentaron información con respecto a *la duración del programa* quedándose fuera un total de 25 documentos sin analizar. Se observa en la tabla 90 las mayores magnitudes del efecto moderado en programas de menor duración –a excepción de las sesiones– como es el caso de los 9 documentos en donde se ha aplicado el aprendizaje cooperativo entre 1 y 4 semanas ( $d=0,79$ ) y en los otros 17 estudios cuando se lleva a cabo entre 5 y 8 semanas ( $d=0,72$ ) aunque también es cierto

que en ambos subgrupos el porcentaje de heterogeneidad es alto (81% y 76%, respectivamente).

En la medida en que avanzan las semanas del semestre, a partir de la novena, los efectos giran en torno al 0,5. Así, el subgrupo entre 9 y 12 semanas y entre 13 y 16 semanas obtienen una magnitud del efecto de 0,5 y 0,51 respectivamente y, a partir de 17 semanas o mas, con un número bajo de documentos ( $n=3$ ) y una  $d=0,55$ .

Las sesiones de dos o tres intervenciones son las que reportan el efecto y la heterogeneidad mas baja ( $d=0,3$ ;  $I^2=19\%$ ).

Tabla 90: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: duración del programa

Duración del programa									
Subgrupos	N° estudios	N° resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Sesiones	15	26	0,3	0,18	0,42	4,91	$p<0,00001$	30,88	19%
Entre 1 y 4 semanas	9	10	0,79	0,64	0,94	10,14	$p<0,00001$	46,41	81%
Entre 5 y 8 semanas	17	19	0,72	0,62	0,81	14,32	$p<0,00001$	74,71	76%
Entre 9 y 12 semanas	14	17	0,5	0,38	0,62	8,34	$p<0,00001$	47,10	66%
Entre 13 y 16 semanas	7	8	0,51	0,32	0,7	5,36	$p<0,00001$	52,07	87%
17 semanas o más	3	3	0,55	0,22	0,87	3,28	$p=0,001$	6,20	68%
Total	65	83							

Fuente: elaboración propia

Al comparar los estudios en función de la prueba de homogeneidad se han encontrado diferencias estadísticamente significativas entre grupos  $Q_b=36,84$ ;  $p<0,0001$  e intragrupos  $Q_w=257,36$ ;  $p<0,0001$ , de forma que es necesario estudiar otras fuentes de heterogeneidad (ver tabla 91).

Tabla 91: Resultados del ANOVA variable moderadora: duración del programa

ANOVA		$Q$	P-value
Duración del programa	Between groups ( $Q_b$ )	36,84	$p<0,0001$
	Within groups ( $Q_w$ )	257,36	$p<0,0001$
	Total	294,19	$p<0,0001$

Fuente: elaboración propia

## Conformación de grupos

La *conformación de los grupos* según los distintos niveles de habilidades de los integrantes, en homogéneos u heterogéneos, es una variable ecológica que solo se especifica en 48 de los 90 documentos incluidos (ver tabla 92).

Los grupos heterogéneos obtienen una magnitud del efecto moderada ( $d=0,56$ ) y los grupos homogéneos un efecto pequeño ( $d=0,41$ ) pero en este último el porcentaje de heterogeneidad es 0% en contraste con los grupos heterogéneos con un porcentaje moderado ( $I^2=67\%$ ).

La prueba de homogeneidad intragrupos es estadísticamente significativa,  $Q_w=137,67$ ;  $p<0,0001$  al igual que la intergrupos  $Q_b=8,73$ ;  $p=0,0127$  de forma que es necesario seguir estudiando otras fuentes de heterogeneidad (ver tabla 93).

Tabla 92: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: conformación de los grupos

Conformación de los grupos									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Homogéneo	16	27	0,41	0,28	0,53	6,45	$p<0,00001$	17,79	0%
Heterogéneo	32	40	0,56	0,48	0,64	13,18	$p<0,00001$	119,88	67%
Total	48	67							

Fuente: elaboración propia

Tabla 93: Resultados del ANOVA variable moderadora: conformación de los grupos

ANOVA		$Q$	P-value
Conformación de los grupos	Between groups ( $Q_b$ )	8,73	0,0127
	Within groups ( $Q_w$ )	137,67	$p<0,0001$
	Total	146,4	$p<0,0001$

Fuente: elaboración propia

## Número de integrantes

Continuando con el análisis de las variables moderadoras, al comparar los resultados de los 68 documentos en los cuales los grupos cooperativos han trabajado *por parejas*, *tríos* y *entre 4 y 6 integrantes* –siendo estas agrupaciones no las únicas pero sí las mas comunes–, se han encontrado diferencias estadísticamente significativas entre esos tres subgrupos,  $Q_b=10,85$ ;  $p=0,0044$  e intragrupos  $Q_w=318,89$ ;  $p<0,0001$ , de forma que es necesario estudiar otras fuentes de heterogeneidad (ver tabla 95).

La magnitud del efecto para grupos conformados por parejas y tríos es similar pero baja ( $d=0,45$  y  $0,43$ , respectivamente) mientras que el trabajo en grupo entre cuatro y seis

integrantes obtiene un tamaño del efecto mayor con una  $d=0,63$  aunque este subgrupo presenta el porcentaje de heterogeneidad mas alto con un 78% (ver tabla 94).

Tabla 94: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: N° de integrantes

N° integrantes									
Subgrupos	N° estudios	N° resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Dos miembros	29	43	0,45	0,37	0,52	11,81	$p<0,00001$	138,59	70%
Tres miembros	8	12	0,43	0,26	0,61	4,89	$p<0,00001$	18,09	39%
Entre cuatro y seis miembros	31	36	0,63	0,55	0,72	14,26	$p<0,00001$	162,21	78%
Total	68*	91							

\* aunque la sumatoria es 68 son realmente 67 documentos porque el estudio de O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986) que tiene dos *outcomes* aparece repetido en el subgrupo “dos miembros” y “tres miembros”

Fuente: elaboración propia

Tabla 95: Resultados del ANOVA variable moderadora: N° de integrantes

ANOVA		$Q$	P-value
N° integrantes	Between groups ( $Q_b$ )	10,85	$p=0,0044$
	Within groups ( $Q_w$ )	318,89	$p<0,0001$
	Total	329,74	$p<0,0001$

Fuente: elaboración propia

Al comparar la eficacia de los grupos cuando el *número de integrantes es par* frente a grupos con un *número de miembros impar*, solo un total de 62 documentos cumplen con este criterio. El trabajo con grupos pares obtuvo un efecto bajo-moderado ( $d=0,49$ ) mientras que los grupos impares un efecto moderado ( $d=0,55$ ) (ver tabla 96). La variabilidad *between groups* revela que no se encontraron diferencias estadísticamente significativas entre ambos subgrupos,  $Q_b=0,68$ ;  $p=0,4096$  pero no así en la variabilidad *within groups*  $Q_w=226,08$ ;  $p<0,0001$  (ver tabla 97).



Tabla 96: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: grupos pares e impares

Grupos pares e impares									
Subgrupos	N° estudios	N° resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Grupos pares	46*	64	0,49	0,43	0,55	15,32	p<0,0001	216,14	71%
Grupos impares	16	20	0,55	0,42	0,67	8,57	p=0,0001	49,94	62%
Total	62	84							

\* en esta clasificación se encuentra solo uno de los tres *outcomes* del estudio de O'Donnell, Dansereau, Hythecker, Larson, Rocklin, Lambiotte & Young (1986)

Fuente: elaboración propia

Tabla 97: Resultados del ANOVA variable moderadora: grupos pares e impares

ANOVA		$Q$	P-value
Grupos pares e impares	Between groups ( $Q_b$ )	0,68	0,4096
	Within groups ( $Q_w$ )	226,08	p<0,0001
	Total	266,76	p<0,0001

Fuente: elaboración propia

### Estructura de recompensa

En un principio, 17 documentos de un total de 90 ofrecieron información con relación a la *estructura de recompensa* pero finalmente se han analizado 15 debido a dos motivos. El primero, se ha excluido del análisis el subgrupo “recompensa de grupo basada en el producto grupal” porque aunque este tipo de recompensa forma parte del sistema de evaluación y el estudio de Bayraktar (2011) bajo el método LT cumple con esta característica, este documento primario solo cuenta con un *outcome*. El segundo, el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado a la vez en mas de un subgrupo –recompensas de grupo basadas en los aprendizajes individuales y sin recompensas de grupo–.

En consecuencia, de un total de 15 documentos, en 4 estudios conformados por métodos cooperativos que no reciben recompensas de grupo –GI y *Jigsaw I*– la magnitud del efecto es moderado con una  $d=0,61$  y una heterogeneidad de 0% frente a aquellos cuyas estructuras exigen recompensas grupales basadas en los aprendizajes dentro del grupo –STAD, TAI y *Jigsaw II*– los cuales obtuvieron un efecto también moderado de  $d=0,55$  pero con una heterogeneidad muy alta (81%) (ver tabla 98).

Tabla 98: Resultados magnitud del efecto y pruebas de homogeneidad variable moderadora: estructura de recompensas

Estructura de recompensa									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Pruebas homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Recompensa de grupo basada en aprendizajes individuales*	12	15	0,55	0,41	0,68	7,77	$p<0,0001$	72,45	81%
Sin recompensa de grupo**	4	4	0,61	0,34	0,89	4,38	$p<0,0001$	1,61	0%
Total	16***	19							

\* incluye los métodos STAD, TGT, TAI, CIRC y *Jigsaw II* pero en el presente estudio no hay métodos TGT ni CIRC

\*\* incluye los métodos GI y *Jigsaw I*

\*\*\* suman 16 documentos pero son 15 porque el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado a la vez en el subgrupo “recompensa de grupo basada en aprendizajes individuales” (un *outcome*) y “sin recompensa grupal” (un *outcome*)

Fuente: elaboración propia

No obstante, las diferencias entre los dos subgrupos no han sido estadísticamente significativas  $Q_b=0,78$ ;  $p=0,678$  pero la variabilidad intraestudios si ha sido estadísticamente significativa  $Q_w=74,06$ ;  $p<0,0001$  dato que indica la necesidad de seguir buscando otras fuentes de variabilidad (ver tabla 99).

Tabla 99: Resultados del ANOVA variable moderadora: estructura de recompensas

ANOVA		$Q$	P-value
Estructura de recompensa	Between groups ( $Q_b$ )	0,78	0,678
	Within groups ( $Q_w$ )	74,06	$p<0,0001$
	Total	74,8	$p<0,0001$

Fuente: elaboración propia

### Estructura de tarea

Otra variable ecológica que ha sido estudiada en los métodos cooperativos es la *estructura de la tarea*, es decir, si cada uno de los integrantes del grupo se especializa en una tarea, rol o función creando así una interdependencia alta entre sus miembros o de lo contrario, no existe una división del material lo que hace que la relación de dependencia entre sus integrantes sea baja.

Solo 17 estudios de 90 cumplen con esta característica y de estos, solo 16 documentos han sido analizados porque el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado a la vez en dos subgrupos –con especialización y sin especialización–.

A penas en 8 estudios con métodos que exigen la especialización de la tarea la magnitud del efecto ha sido grande con  $d=0,88$  y en 9 estudios cuando los métodos no exigen tal especialización el efecto ha sido débil con una  $d=0,36$ . En el primer subgrupo la

heterogeneidad es alta con un 81% y en el segundo es moderada con 46%. La variabilidad intragrupos es estadísticamente significativa  $Q_w=57,6$ ;  $p<0,0001$  y también la variabilidad entre grupos  $Q_b=17,23$ ;  $p<0,0001$  (ver tablas 100 y 101).

Tabla 100: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: estructura de tarea

Estructura de la tarea									
Subgrupos	N° estudios	N° resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Con especialización*	8	8	0,88	0,69	1,07	9,09	$p<0,00001$	37,10	81%
Sin especialización**	9	12	0,36	0,21	0,52	4,54	$p<0,00001$	20,50	46%
Total	17***	20							

\* incluye los métodos *Jigsaw I*, *Jigsaw II* y GI

\*\* incluye los métodos STAD, TGT, TAI, CIRC y LT pero en el presente estudio no hay métodos TGT ni CIRC

\*\*\* suman 17 documentos pero son 16 porque el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado a la vez en el subgrupo “con especialización” y en el subgrupo “sin especialización”

Fuente: elaboración propia

Tabla 101: Resultados ANOVA variable moderadora: estructura de tarea

ANOVA		$Q$	P-value
Estructura de la tarea	Between groups ( $Q_b$ )	17,23	$p<0,0001$
	Within groups ( $Q_w$ )	57,6	$p<0,0001$
	Total	74,83	$p<0,0001$

Fuente: elaboración propia

### Igualdad de oportunidades en la puntuación

Por último, la *igualdad de oportunidades en la puntuación* refiere a que si cada uno de los integrantes del grupo de manera individual se esfuerza al máximo podrá contribuir a la puntuaciones de su equipo o si por el contrario, las puntuaciones son diferentes para cada estudiante.

Por tratarse de una variable que al igual que la estructura de recompensas y la estructura de tareas depende del método cooperativo empleado para su clasificación, solo 17 estudios de 90 han cumplido con esta condición, de los cuales se han analizado 16 porque se repite el mismo caso que en la estructura de la tarea: el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado cada uno a la vez en los dos subgrupos.

Este subgrupo no ha obtenido una significación estadística entre los subgrupos  $Q_b=0,469$ ;  $p=0,494$  pero si intrasubgrupos con una  $Q_w=74,36$ ;  $p<0,0001$  (ver tabla 102)

Tabla 102: Resultados del ANOVA variable moderadora: igualdad oportunidades puntuación

ANOVA		$Q$	P-value
Igualdad de oportunidades en la puntuación	Between groups ( $Q_b$ )	0,469	p=0,494
	Within groups ( $Q_w$ )	74,36	p<0,0001
	Total	74,83	p<0,0001

Fuente: elaboración propia

Se observa una magnitud moderada del efecto ( $d=0,55$ ) en los 12 documentos en donde los métodos cooperativos aplicados favorecen la igualdad de oportunidades en la puntuación (con 81% de heterogeneidad) y también en aquellos 5 en donde el esfuerzo individual no aporta a la calificación grupal ( $d=0,65$ ), este último con un 0% de heterogeneidad (ver tabla 103).

Tabla 103: Resultados magnitud del efecto y pruebas homogeneidad variable moderadora: igualdad oportunidades puntuación

Igualdad de oportunidades en la puntuación									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Con igualdad*	12	15	0,55	0,41	0,68	7,77	p<0,00001	72,45	81%
Sin igualdad**	5	5	0,65	0,40	0,89	5,12	p<0,00001	1,92	0%
Total	17***	20							

\* incluye los métodos STAD, TGT, TAI, CIRC y *Jigsaw II* pero en el presente estudio no hay métodos TGT ni CIRC\*\* incluye los métodos LT, *Jigsaw I* y GI\*\*\* suman 17 documentos pero son 16 porque el estudio de Jalilifar (2010) que reporta dos *outcomes* ha sido clasificado a la vez en el subgrupo "con igualdad" (un *outcome*) y en el subgrupo "sin igualdad" (un *outcome*)

Fuente: elaboración propia

#### 5.1.2.5. Análisis por estratos

Un análisis por estratos ha permitido estudiar las áreas de conocimiento en función de los métodos cooperativos empleados en *Ciencias*, *Ciencias Sociales y Humanidades* y *Artes*, no así en *Ingeniería*, *Industria y Construcción*; *Salud y Servicios Sociales* y *Educación*. El motivo obedece a que en estas tres últimas áreas o bien, solo se cuenta con un resultado por grupo lo que imposibilita el cálculo de la magnitud del efecto o bien, los métodos cooperativos no cumplen con los criterios de clasificación.

Por ejemplo, en el caso de *Ingeniería*, *Industria y Construcción* que cuenta con tres estudios, en dos de ellos no se ha especificado el método (Hsiung, 2011; Hsiung 2012) y el tercero pertenece a la clasificación de *Computer Support* pero con un solo *outcome* (Yaeger, Marra, Costanzo & Gray, 1999).

Un comportamiento similar ha ocurrido con los tres documentos del área de *Salud y Servicios Sociales*. El estudio de Baghcheghi, Reza & Rezaei (2011) no pertenece a

ninguna clasificación de los métodos cooperativos y los otros dos aunque conciernen a *Dyadic Methods* (Iwasiw & Goldenberg, 1993) y *Computer Support* (Lin, 2012) no cuentan con al menos un *outcome* para el cálculo del efecto.

Por último, pese a que *Educación* dispone del mayor número de documentos, 6 en total, 4 de ellos no se pueden clasificar según el método cooperativo empleado (Hwong, Caswell, Johnson & Johnson, 1993; Jimison, 1990; Kiliç, 2008; Maden, 2010) y los otros dos solo cuentan con un *outcome* cada uno (Bayraktar, 2011; Perihan & Kamuran 2007).

#### *Ciencias + métodos cooperativos*

Un total de 19 documentos han sido agrupados bajo el área de conocimiento *Ciencias* y que a la vez han logrado clasificarse según los *métodos cooperativos* empleados pero se ha excluido el estudio de Courtney, Courtney y Nicholson (1992) debido a que aunque trabaja con Ciencias y el método es *Informal Methods* no dispone de al menos dos *outcomes*. De esta forma se ha reducido a 18 el número definitivo de estudios a analizar.

En la tabla 104 se observa como en todos los subgrupos se han encontrado diferencias significativas a favor del grupo experimental pero no así cuando se trabajan asignaturas del área de *Ciencias* con métodos cooperativos en STL en donde la magnitud del efecto ha sido positiva pero débil ( $d= 0,15$ ) y a la vez no significativa ( $IC_{95\%}= -0,07; +0,37$ ;  $p=0,169$ ).

Por tanto, cuando las diferencias son a favor del grupo experimental, la magnitud del efecto es grande en asignaturas del área de *Ciencias* con métodos cooperativos *combinados* ( $d=0,83$ ;  $n=2$ ) pero con una heterogeneidad del 92% y es moderada, cuando en las *Ciencias* se emplean métodos cooperativos que se agrupan bajo las clasificaciones *Cooperative Investigation* ( $d=0,72$ ;  $n=2$ ) con un índice  $I^2$  de 0% y *Computer Support* ( $d=0,74$ ;  $n=12$ ) en donde la heterogeneidad es alta con un 76%

Tabla 104: Resultados magnitud del efecto y pruebas homogeneidad estrato: Ciencias + métodos cooperativos

Área de conocimiento: Ciencias + métodos cooperativos									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Ciencias + Cooperative Investigation	2	2	0,72	0,31	1,14	3,42	$p=0,0006$	0,001	0%
Ciencias + combinados	2	2	0,83	0,54	1,12	5,64	$p<0,00001$	12,41	92%
Ciencias + Computer Support	9	12	0,74	0,58	0,90	9,18	$p<0,00001$	46,1	76%
Ciencias + STL	5	7	0,15	-0,07	0,37	1,38	$p=0,169$	6,13	2%
Total	18	23							

Fuente: elaboración propia

Sin embargo, el análisis ANOVA indica que existen diferencias significativas intersubgrupos  $Q_b=21,9$ ;  $p=0,0001$  e intrasubgrupos  $Q_w=64,65$ ;  $p<0,0001$ , datos que invitan a continuar indagando posibles fuentes de heterogeneidad debido a que esa variabilidad no se debe al azar (ver tabla 105).

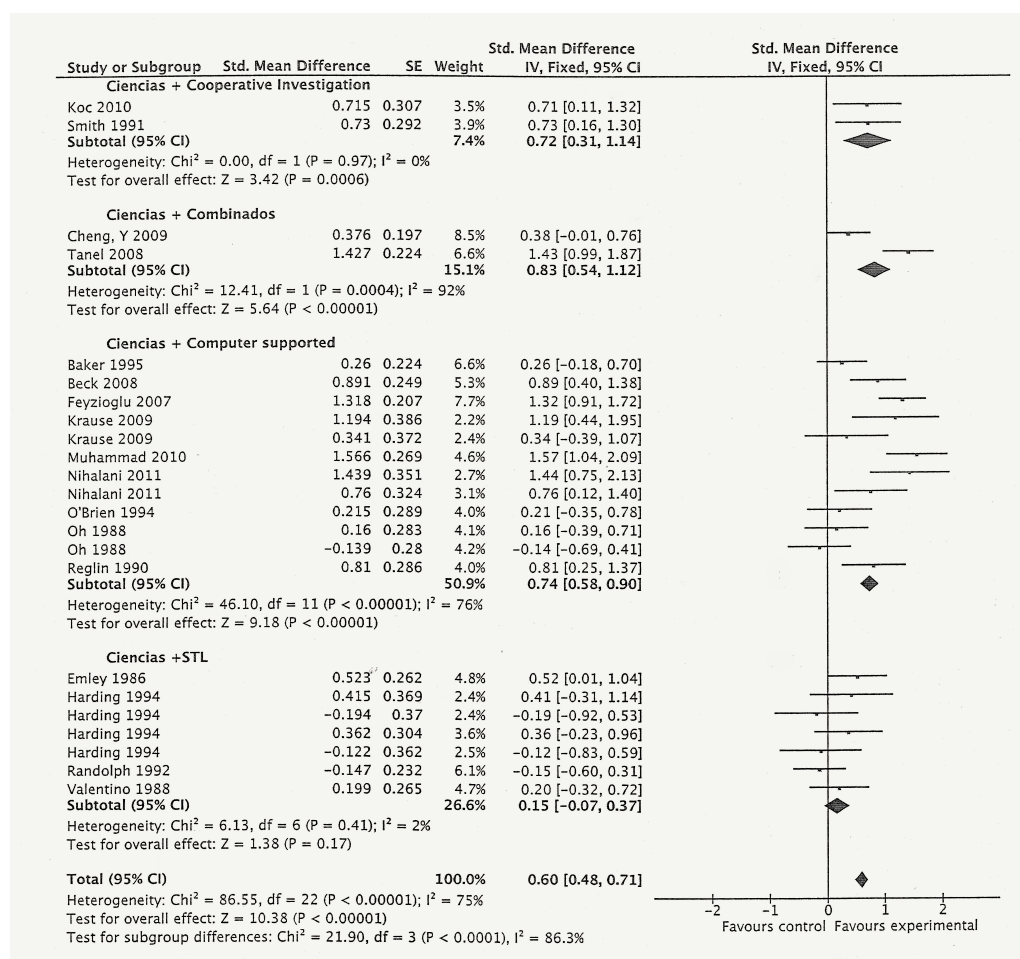
Tabla 105: Resultados ANOVA estrato: Ciencias + métodos cooperativos

ANOVA		$Q$	P-value
Ciencias + métodos cooperativos	Between groups ( $Q_b$ )	21,9	$p=0,0001$
	Within groups ( $Q_w$ )	64,65	$p<0,0001$
	Total	86,55	$p<0,0001$

Fuente: elaboración propia

El gráfico de efectos o *forest plot* representa y resume gráficamente el análisis de lo ocurrido en este estrato, distinguiéndose así los estudios individuales con efectos positivos de aquellos con efectos negativos. También indica los que han obtenido un efecto significativo porque el intervalo de confianza no cruza la línea vertical, es decir, se encuentra por completo en una de las dos zonas (ver ilustración 11):

Ilustración 11: Florest plot magnitud del efecto estrato: Ciencias + métodos cooperativos



Fuente: elaboración propia

*Ciencias Sociales + métodos cooperativos*

De los 15 documentos que han cumplido con ambos criterios solo 9 se han utilizado para el cálculo de la magnitud del efecto porque existen tres estudios que se han excluido porque cuentan cada uno con un solo *outcome*, los cuales se han trabajado en *Ciencias Sociales* con métodos cooperativos bajo la clasificación STL (Van Wyk, 2010), *Informal Methods* (Klein & Schnackenberg, 2000) y *Cooperative Investigation* (Tran & Lewis, 2012) y además, los restantes 3 documentos no fueron clasificados bajo ningún método cooperativo(ver tabla 106).

La magnitud del efecto es positiva pero débil y significativa a favor del grupo experimental cuando se trabaja en *Ciencias Sociales* con métodos cooperativos combinados ( $d=0,43$ ;  $n=2$ ) y con *Computer Support* ( $d=0,25$ ;  $n=4$ ) y moderada bajo la

denominación de *Dyadic Methods* ( $d=0,67$ ;  $n=3$ ). En estos dos últimos subgrupos la heterogeneidad es de 0%, lo que significa que esta variabilidad se debe al azar y no a otras variables moderadoras. Tampoco en esta variable se han encontrado diferencias significativas entre subgrupos  $Q_b=5,07$ ;  $p=0,0793$  e intrasubgrupos  $Q_w=10,68$ ;  $p=0,2203$  (ver tabla 107).

Tabla 106: Resultados magnitud del efecto y pruebas homogeneidad estrato: Ciencias Sociales + métodos cooperativos

Área de conocimiento: Ciencias Sociales + métodos cooperativos									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Ciencias Sociales + Dyadic Methods	3	5	0,67	0,37	0,97	4,44	$p<0,00001$	9,78	59%
Ciencias Sociales + combinados	2	2	0,43	0,18	0,68	3,39	$p=0,0007$	0,37	0%
Ciencias Sociales + Computer Support	4	4	0,25	0,04	0,47	2,31	$p=0,02$	0,54	0%
Total	9	11							

Fuente: elaboración propia

Tabla 107: Resultados ANOVA estrato: Ciencias Sociales + métodos cooperativos

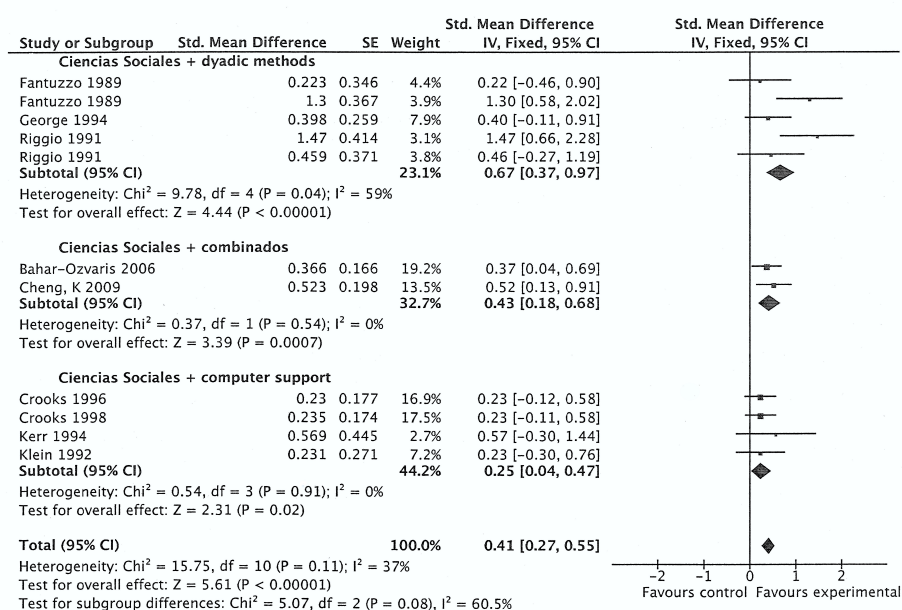
ANOVA		$Q$	P-value
Ciencias Sociales + métodos cooperativos	Between groups ( $Q_b$ )	5,07	$p=0,0793$
	Within groups ( $Q_w$ )	10,68	$p=0,2203$
	Total	15,75	$p=0,1069$

Fuente: elaboración propia

A manera ilustrativa, el gráfico *forest plot* permite nuevamente ver las magnitudes del efecto medio de cada estudio pero de una manera mucho mas esclarecedora (ver ilustración 12):



Ilustración 12: Florest plot magnitud del efecto estrato: Ciencias Sociales + métodos cooperativos



Fuente: elaboración propia

### Humanidades y Artes + métodos cooperativos

De los 11 estudios que han cumplido con las características de este estrato solo se han seleccionado un total de 8. Los tres estudios restantes se han descartado porque aunque existen dos estudios en donde se han aplicado métodos pertenecientes a las clasificaciones *Informal Methods* (Boling & Robinson, 1999) y *combinados* (Karababa, 2009) cada uno presenta solo un *outcome*. El tercer resultado excluido ha sido uno de los dos *outcomes* del estudio de Jalilifar (2010), el que pertenece a la clasificación *Cooperative Investigation* no así el otro, que se ha clasificado bajo STL.

Por consiguiente, el análisis se ha reducido a dos subgrupos, ambos significativos a favor del grupo experimental. El primero, con solo dos estudios en donde la magnitud del efecto es alta, con una de  $d=1,06$  y una heterogeneidad de  $0\%$  cuando en asignaturas de *Humanidades y Artes* se aplican métodos cooperativos *Computer Support*. El segundo, con seis estudios y con un efecto moderado de  $0,69$  en donde se ha utilizado en esta área de conocimiento los métodos agrupados en STL pero con una heterogeneidad de  $88\%$  y con a penas 2 estudios (ver tabla 108).

Tabla 108: Resultados magnitud del efecto y pruebas homogeneidad estrato: Humanidades/Artes + métodos cooperativos

Área de conocimiento: Humanidades/Artes + métodos cooperativos									
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad	
				LI	LS	Z-value	P-value	$Q$	$I^2$
Humanidades /Artes + STL	6	6	0,69	0,47	0,91	6,11	$p<0,00001$	41,15	88%
Humanidades /Artes + Computer Supported	2	2	1,06	0,69	1,42	5,70	$p<0,00001$	0,86	0%
Total	8	8							

Fuente: elaboración propia

El análisis *between groups* indica que no existen diferencias significativas entre los subgrupos  $Q_b=2,93$ ;  $p=0,0871$  pero el análisis *within groups* si indica que existen diferencias significativas intragrupos  $Q_w=42,02$ ;  $p<0,0001$ . Otra vez mas resulta necesario continuar indagando otras variables explicativas (ver tabla 109).

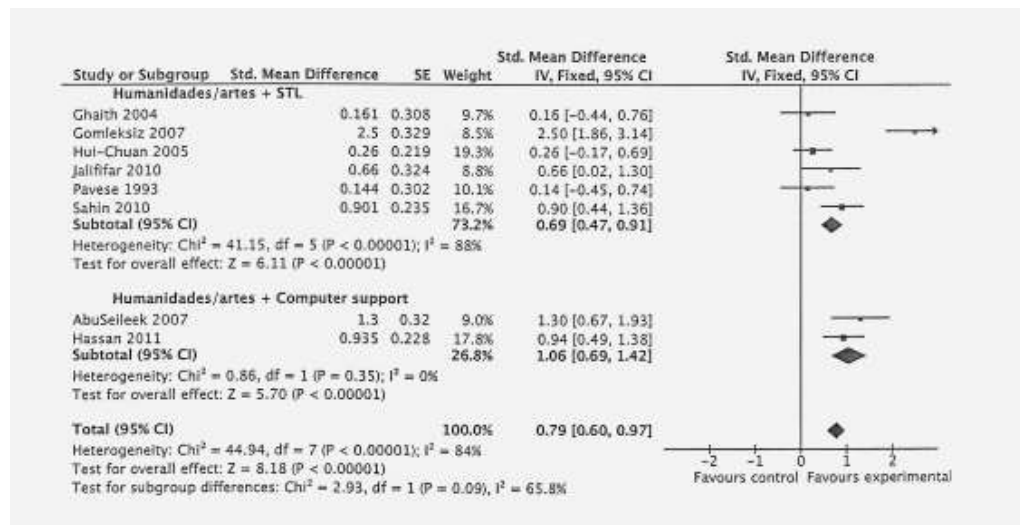
Tabla 109: Resultados ANOVA estrato: Humanidades/Artes + métodos cooperativos

ANOVA		$Q$	P-value
Humanidades/Artes + métodos cooperativos	Between groups ( $Q_b$ )	2,93	$p=0,0871$
	Within groups ( $Q_w$ )	42,02	$p<0,0001$
	Total	44,94	$p<0,0001$

Fuente: elaboración propia

A continuación el gráfico *forest plot* representa los resultados de los estudios individuales en donde se evidencia el valor medio de cada uno siendo una técnica de difusión muy efectiva para obtener en una sola visión una gran cantidad de información (ver ilustración 13):

Ilustración 13: Florest plot magnitud del efecto estrato: Humanidades/Artes + métodos cooperativos

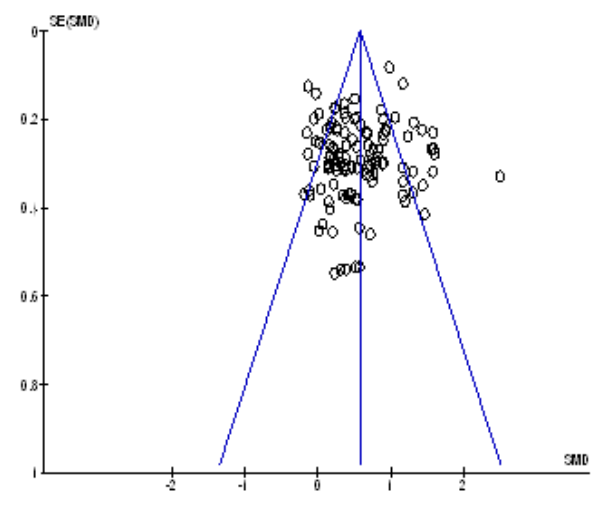


Fuente: elaboración propia

#### 5.1.2.6. Sesgo de publicación

En el gráfico de embudo o *funnel plot*, aparentemente, no se aprecia asimetría en la distribución de los puntos, por lo que partir de la observación de este gráfico se rechazaría la presencia de sesgo de publicación (ver ilustración 14).

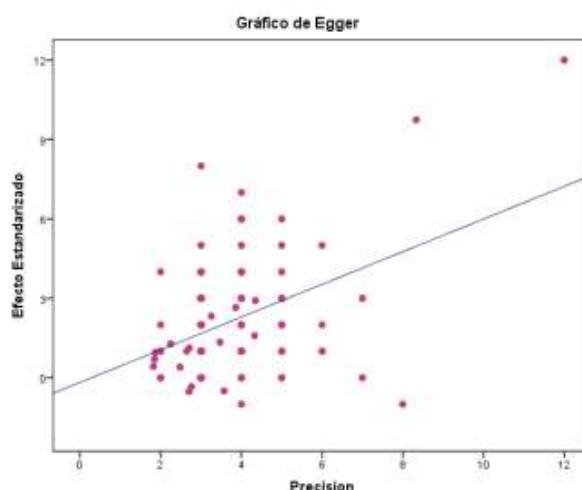
Ilustración 14: Funnel plot: sesgo de publicación



Fuente: elaboración propia

Este resultado se confirma también en el test de Egger,  $p=0,727$  (Egger, Smith, Schneider & Minder, 1997) donde tampoco se tienen evidencias para rechazar la hipótesis nula sobre el término independiente (el coeficiente es igual a 0) resultado que ratifica nuevamente la ausencia de sesgo de publicación (ver ilustración 15).

Ilustración 15: Gráfico de Egger: sesgo de publicación



Fuente: elaboración propia

Otro análisis que se ha realizado es el cálculo del número de seguridad de Orwin (1983) que es una variante del método de Rosenthal (1979).

Para su cálculo se ha considerado que la magnitud del efecto medio para los estudios no recuperados es 0 y la magnitud del efecto crítico es 0,2.

Los resultados indican que 214,2 es un número muy alto de estudios no recuperados, por lo que la probabilidad de que el sesgo de publicación haya modificado sustancialmente los resultados es baja

$$N_s = \frac{k_r(\bar{d}_r - \bar{d}_c)}{d_c - \bar{d}_a} = \frac{119(0,56 - 0,2)}{0,2 - 0} = 214,2$$

### 5.1.3. Síntesis cualitativa: calidad de los estudios

Estudiar la *calidad metodológica* de los estudios primarios como otra variable moderadora de interés, ha permitido conocer su incidencia en los efectos del aprendizaje cooperativo en el rendimiento de estudiantes universitarios (ver en el anexo 21 el gráfico *forest plot*).

Es llamativo observar cómo en función del nivel de calidad de los estudios el efecto no varía. En estudios con baja calidad la magnitud del efecto es moderada con una  $d=0,74$  y se sigue manteniendo con un nivel medio y alto ( $d=0,52$  y  $0,58$ , respectivamente). También es cierto que las diferencias encontradas son significativas a favor del grupo experimental y que en los tres subgrupos la heterogeneidad es moderada (ver tabla 110).

Si la magnitud global del efecto en estudios sin calidad ha sido moderada con  $0,56$  (modelo de efectos fijos) este tamaño se conserva igualmente en estudios con baja, media y alta calidad aunque el nivel bajo ha obtenido el efecto mas alto de los tres ( $d=0,74$ ).

Tabla 110: Resultados magnitud del efecto y pruebas homogeneidad estrato: calidad metodológica

Calidad metodológica										Sin calidad
Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad		$d$ 0,56
				LI	LS	Z-value	P-value	$Q$	$I^2$	
Baja	10	14	0,74	0,63	0,86	12,64	$p<0,00001$	28,78	55%	
Media	52	72	0,52	0,47	0,58	17,60	$p<0,00001$	254,09	72%	
Alta	28	33	0,58	0,49	0,67	12,84	$p<0,00001$	136,55	77%	
Total	90	119								

Fuente: elaboración propia

Las pruebas de variabilidad entre grupos  $Q_b=11,02$ ;  $p=0,004$  y variabilidad intragrupos  $Q_w=419,42$ ;  $p<0,0001$  han sido significativas, valores que indican que existen otras fuentes explicativas del tamaño de los efectos encontrados (ver tabla 111).

Tabla 111: Resultados del ANOVA variable moderadora: calidad metodológica

ANOVA		$Q$	P-value
Calidad metodológica	Between groups ( $Q_b$ )	11,02	0,004
	Within groups ( $Q_w$ )	419,42	p<0,0001
	$Q_{total}$	430,44	p<0,0001

Fuente: elaboración propia

Explorar a continuación los efectos de la calidad según los *años de publicación*, *métodos cooperativos* y las *áreas de conocimiento*, es profundizar en algunos subgrupos de variables extrínsecas y ecológicas que ya apuntaban en análisis previos con magnitudes del efecto moderadas–altas.

A primera vista se observa en la tabla 112 que en ninguno de los tres subgrupos en el nivel de calidad baja se ha logrado calcular la magnitud del efecto porque no se cuenta con suficientes *outcomes* o no existen grupos que puedan ser categorizados. En los otros dos niveles de calidad restantes, todos ofrecen información menos los métodos cooperativos bajo la clasificación de *Cooperative Investigation e Informal Methods* y, el área de Salud y Servicios Sociales.

A esta realidad se suma, en algunos subgrupos un bajo número de documentos o ninguno que lo detenta, por lo que los resultados obtenidos en los siguientes análisis tendrán que asumirse con precaución. Tal es el caso del área de Ingeniería en donde no ha sido posible valorar su eficacia según la calidad metodológica en ninguno de los estratos estudiados.

Tabla 112: Resumen magnitud del efecto estrato: calidad metodológica + años de publicación + métodos cooperativos + áreas de conocimiento

Calidad metodológica						
	Baja <i>d</i>	Nº	Media* <i>d</i>	Nº	Alta <i>d</i>	Nº
<b>Años de publicación</b>						
≤ 1999			0,30	31	0,36	9
≥ 2000			0,77	21	0,69	19
<b>Métodos cooperativos</b>						
STL			0,23	4	0,73	8
Cooperative Investigation					0,58	3
Dyadic Methods			0,43	12	0,47	1
Informal Methods					0,49	2
Computer Support			0,72	12	0,58	4
Combinados			0,74	2	0,26	2
<b>Áreas de conocimiento</b>						
Ciencias			0,63	23	0,82	7
Ciencias Sociales, Educación Comercial y Derecho			0,24	9	0,47	5
Ingeniería, Industria y Construcción						
Humanidades y Artes			0,60	6	0,49	10
Salud y Servicios Sociales			1,06	2		
Educación			0,78	3	0,64	3

Fuente: elaboración propia

Continuando con el análisis general de la tabla anterior (112) se observa que independientemente del nivel de calidad solo mantienen la misma magnitud del efecto los años de publicación entre 1980-1999 y 2000-2012, los métodos cooperativos por parejas *Dyadic Methods*, los que se apoyan con tecnología *Computer Support* y las áreas de Ciencias Sociales y Educación (ver tabla 112).

Los métodos cooperativos combinados y las Humanidades y Artes disminuyen sus tamaños del efecto cuando aumentan los niveles de calidad pero el método STL y las Ciencias presentan comportamientos diferentes porque aumentan sus magnitudes en la medida en que también se incrementan sus niveles de calidad.

Conocida estas estimaciones, en las próximas páginas se presenta un análisis por estratos de los años de publicación, los métodos cooperativos y las áreas de conocimiento comparando entre estudios *con* y *sin calidad metodológica* para luego al final del apartado mostrar una tabla resumen de la variabilidad entre grupos e intragrupos (estadístico *Q*).

*Años de publicación + Con y sin calidad metodológica*

El nivel bajo de calidad ha sido imposible de calcular porque entre los años 2000-2012 hay un solo el estudio que es el de Cheng (2009) lo que no permite a su vez la comparación con estudios entre los años 1980-1999 que si cuenta con 9 estudios (ver tabla 113).

Es interesante observar que la magnitud del efecto en los años de publicación no cambia en estudios con y sin calidad metodológica. Por ejemplo, los estudios con calidad media y alta entre los años 1980-1999 han obtenido un tamaño pequeño del efecto ( $d=0,30$  y  $0,36$ , respectivamente) y esta estimación se mantiene con estudios sin calidad ( $d=0,43$ ). Lo mismo acontece entre los años 2000-2012 donde la magnitud ha sido mediana para estudios con calidad media ( $d=0,77$ ) y alta ( $d=0,69$ ) y se conserva con una  $d=0,73$  en estudios sin calidad.

Tabla 113: Resultados magnitud del efecto estrato: calidad metodológica + años de publicación

Calidad metodológica + años de publicación											Sin calidad
	Subgrupos	Nº estudios	Nº resultados independientes	$d$	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad		$d$
					LI	LS	Z-value	P-value	$Q$	$I^2$	
Media	Años ≤1999	31	48	0,30	0,22	0,38	7,15	$p<0,00001$	52,36	10%	0,43
	Años ≥2000	21	24	0,77	0,68	0,85	17,92	$p<0,00001$	139,19	83%	0,73
Alta	Años ≤1999	9	11	0,36	0,21	0,51	4,68	$p<0,00001$	18,39	46%	0,43
	Años ≥2000	19	22	0,69	0,58	0,80	12,44	$p<0,00001$	106,32	80%	0,73

Fuente: elaboración propia

*Métodos cooperativos + Con y sin calidad metodológica*

En este estrato, tampoco se ha podido calcular el nivel bajo de calidad metodológica porque los métodos cooperativos *Cooperative Investigation*, métodos combinados y *Computer Support* reportan un solo *outcome* por subgrupo. Son Smith, Hinckley y Volk (1991), Cheng (2009) y Yaeger, Marra, Contanzo y Gray (1999) los que quedan fuera para el cálculo no así el estudio de O'Donnell, Dansereau, Rocklin, Hythecker, Lambiotte, Larson y Young (1985) que aunque reporta dos *outcomes* no tendría otro subgrupo para contrastarse.



A esta dificultad se suman otros subgrupos que tampoco pueden analizarse debido a que solo cuentan con un *outcome*. En el nivel de calidad media ocurre con el estudio de Bayraktar (2011) para el método LT y Courtney, Courtney & Nicholson (1992) para *Informal Methods*. En el nivel de calidad alta solo con el estudio de Jalilifar (2010) para el método STL. Asimismo *Cooperative Investigation* e *Informal Methods* no pueden analizarse debido a que aunque si cuentan con mas de un *outcome* solo ofrece información en el nivel alto.

En cambio, de los que si se dispone de información en al menos dos niveles de calidad solo *Dyadic Methods* y *Computer Support* no varían sus magnitudes en estudios con y sin calidad metodológica. El primero, con un efecto pequeño sin calidad ( $d=0,42$ ) y en los niveles medio ( $d=0,43$ ) y alto ( $d=0,47$ ). El segundo, con un tamaño moderado sin calidad ( $d=0,67$ ) que se mantiene de la misma forma en los niveles medio ( $d=0,72$ ) y alto ( $d=0,58$ ) (ver tabla 114).

En contraste, cuando el método es combinado la magnitud es moderada sin calidad ( $d=0,53$ ) y con un nivel medio ( $d=0,74$ ) pero no así con un nivel alto porque disminuye considerablemente a un tamaño pequeño ( $d=0,26$ ) no siendo significativo ( $IC_{95\%} = -0,03; +0,55; p=0,07$ ).

El STL también se modifica. Sin calidad metodológica obtiene un tamaño moderado ( $d=0,55$ ) que se mantiene solo en estudios con alta calidad ( $d=0,73$ ) pero que desciende a un efecto pequeño ( $d=0,23$ ) con estudios de un nivel medio aunque estadísticamente no es significativo ( $IC_{95\%} = -0,00; +0,46; p=0,05$ ).

Tabla 114: Resultados magnitud del efecto estrato: calidad metodológica + métodos cooperativos

Calidad metodológica + métodos cooperativos											Sin calidad
	Subgrupos	Nº estudios	Nº resultados independientes	<i>d</i>	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad		<i>d</i>
					LI	LS	Z-value	P-value	<i>Q</i>	<i>I</i> <sup>2</sup>	
Media	STL	4	7	0,23	-0,00	0,46	1,95	p=0,05	4,04	0%	0,55
	Dyadic Methods	12	20	0,43	0,28	0,58	5,58	p<0,00001	17,78	0%	0,42
	Combinados	2	2	0,74	0,48	1,00	5,56	p<0,00001	14,48	93%	0,53
	Computer Support	12	14	0,72	0,59	0,86	10,48	p<0,00001	59,22	78%	0,67
Alta	STL	8	8	0,73	0,55	0,90	8,26	p<0,00001	56,86	88%	0,55
	Cooperative Investigation	3	3	0,58	0,27	0,89	3,63	p=0,0003	1,40	0%	0,61
	Dyadic Methods	1	3	0,47	0,09	0,85	2,45	p=0,01	2,48	19%	0,42
	Combinados	2	2	0,26	-0,03	0,55	1,78	p=0,07	0,75	0%	0,53
	Computer Support	4	5	0,58	0,35	0,81	4,99	p<0,00001	21,97	82%	0,67
	Informal Methods	2	2	0,49	0,21	0,77	3,38	p=0,0007	1,07	7%	0,37

Fuente: elaboración propia

*Áreas de conocimiento + Con y sin calidad metodológica*

Es una constante al igual que en los dos estratos anteriores la imposibilidad del cálculo de la magnitud del efecto en el nivel bajo de calidad porque en Ciencias Sociales hay un solo *outcome* el de Cheng (2009) situación que se reitera en Ingeniería con Yaeger, Marra, Contanzo y Gray (1999) aunque en Ciencias si hay mas de un estudio pero no tendría otro subgrupo con quien compararse.

Tampoco puede analizarse en el nivel de calidad media el área de Ingeniería porque solo el estudio de Hsiung (2011) cuenta con un *outcome*. Esta situación aparece nuevamente en el nivel de calidad alto pero con un mayor número de estudios como son Klein y Schnackenberg (2000) en Ciencias Sociales, Hsiung (2012) en Ingeniería, Perihan y Kamuran (2007) en Educación y Baghcheghi, Reza y Rezaei (2011) en Salud.

Las áreas que sí pueden contrastarse son todas las que aparecen en la tabla 115 menos Salud y Servicios Sociales por contar solo información en el nivel medio de calidad.

Del resto, Ciencias Sociales mantiene el mismo efecto pequeño sin calidad metodológica ( $d=0,33$ ) que con un nivel medio ( $d=0,24$ ) y alto de calidad ( $d=0,47$ ). Asimismo Educación permanece sin alterar su tamaño. En un nivel medio y alto de calidad con una magnitud moderada ( $d=0,78$  y  $0,64$ , respectivamente) obtiene un valor muy parecido con estudios sin calidad ( $d=0,7$ ).

No así con Humanidades y Artes que pasa de una magnitud media sin calidad ( $d=0,52$ ) que persiste solo con estudios de calidad media ( $d=0,60$ ) y luego disminuye a un efecto pequeño con estudios de alta calidad ( $d=0,49$ ).

Un cambio de tamaño también se evidencia en las Ciencias cuando solo los estudios sin calidad ( $d=0,71$ ) y con un nivel medio ( $d=0,63$ ) logran un efecto moderado pero con una alta calidad la magnitud se modifica a grande ( $d=0,82$ ).

Tabla 115: Resultados magnitud del efecto estrato: calidad metodológica + áreas de conocimiento

Calidad metodológica + áreas de conocimiento											Sin calid d
	Subgrupos	Nº estudios	Nº resultad os indep dientes	<i>d</i>	IC <sub>95%</sub>		Test of overall effect		Prueba homogeneidad		<i>d</i>
					LI	LS	Z- value	P-value	<i>Q</i>	I <sup>2</sup>	
Media	Ciencias	23	30	0,63	0,55	0,71	15,47	p<0,00001	141,68	80%	0,71
	Ciencias Sociales, Educación comercial y Derecho	9	14	0,24	0,12	0,36	3,82	p=0,0001	30,99	58%	0,33
	Humanidade s y Artes	6	8	0,60	0,38	0,82	5,25	p<0,00001	18,90	63%	0,52
	Salud y Servicios Sociales	2	2	1,06	0,71	1,41	5,90	p<0,00001	12,34	92%	1,08
	Educación	3	3	0,78	0,47	1,09	4,92	p<0,00001	0,83	0%	0,7
Alta	Ciencias	7	9	0,82	0,64	1,01	8,91	p<0,00001	47,90	83%	0,71
	Ciencias Sociales, Educación Comercial y Derecho	5	5	0,47	0,30	0,65	5,19	p<0,00001	8,19	51%	0,33
	Humanidade s y Artes	10	11	0,49	0,33	0,64	6,25	p<0,00001	50,89	80%	0,52
	Educación	3	3	0,64	0,34	0,93	4,25	p<0,0001	12,25	84%	0,7

Fuente: elaboración propia

Finalmente, la tabla 116 resume la variabilidad *between groups* y *within groups* según los estratos estudiados en donde se observa que todos son significativos a excepción del subgrupo con un nivel alto de calidad metodológica en métodos cooperativos que no es estadísticamente significativo entre grupos  $Q_b=8,02$ ;  $p=0,1554$  pero si intragrupos  $Q_w=84,52$ ;  $p=0,0001$ .

Tabla 116: Resumen resultados del ANOVA estratos calidad metodológica

Subgrupos	$Q_{total}$	P-value	$Q_b$	P-value	$Q_w$	P-value
Media + Años de publicación	254,09	p<0,0001	62,54	p<0,0001	191,55	p<0,0001
Alta + Años de publicación	136,55	p<0,0001	11,84	0,0006	124,71	p<0,0001
Media + Métodos cooperativos	113,85	p<0,0001	18,33	0,0004	95,52	p<0,0001
Alta + Métodos cooperativos	92,54	p<0,0001	8,02	0,1554	84,52	p<0,0001
Media + Áreas de conocimiento	244,07	p<0,0001	39,32	p<0,0001	204,74	p<0,0001
Alta + Áreas de conocimiento	129,15	p<0,0001	9,93	0,0192	119,23	p<0,0001

Fuente: elaboración propia

## 5.2. Discusión de resultados

No es reciente afirmar que el aprendizaje individual es el mas frecuente en la universidad, el mas criticado y el mas extendido en el ámbito universitario (Nelson, 1986) pero tampoco es menos cierto que “in general, about 20% of the variation in achievement is accounted for by their participation in classroom learning” (Bloom, 1976, p. 123).

Es en esa participación propia del aprendizaje cooperativo como metodología consolidada (Cohen, 1994a), en que los estudiantes aprenden en lo individual y en lo colectivo, con los otros y para los otros, donde el aval de los últimos 30 años de numerosos estudios empíricos publicados en la literatura científica (Johnson & Johnson, 1989; Slavin, 1995) dan buenas referencias a este tipo de enseñanza en el desarrollo cognitivo, social y emocional.

Sin embargo, la mayoría de las síntesis cuantitativas que han estudiado la efectividad del aprendizaje cooperativo en el rendimiento académico de estudiantes se han centrado en primaria, secundaria y bachillerato (Neber, Finsterwald & Urban, 2001; Sharan, 1980; Slavin, Cheung, Groff & Lake, 2008; Slavin, Lake & Groof, 2009, entre otros) no así en la universidad.

Los escasos meta-análisis que incluyen a la Educación Superior como objeto de estudio para la valoración del rendimiento en situaciones cooperativas e individuales se reducen

a dos: Springer, Stanne y Donovan en el año 1999 y el mas recientemente, del año 2002a realizado por los hermanos Johnson.

En el resto de las síntesis cuantitativas, son contados los casos en que la universidad se ha considerado como variable moderadora (Johnson & Johnson, 1989) porque la mayoría de las investigaciones ha asumido esta etapa educativa como una característica mas de los estudios, con el agravante de obtener algunas de estas síntesis el número mas bajo de documentos primarios cuando se compara con otras etapas educativas: 13,7% en la revisión meta-analítica de Lou, Abrami, Spence, Poulsen, Chambers & D'Apollonia (1996), 23,41% en Johnson, Johnson & Stanne (2000) y 25% en Johnson & Johnson (1989), no así en Stanne, Johnson y Johnson (1999) con un 41,4% y Bowen (2000) con 73,33%, aunque otros ni siquiera especifican esta información (Johnson, Maruyama, Johnson, Nelson & Skon, 1981; Lou, Abrami & D'Apollonia, 2001; Lou, Abrami & Spence, 2000).

Además, los reducidos casos en que se menciona el porcentaje de estudios en la universidad, ofrecen magnitudes del efecto globales que miden la eficacia, o no, del aprendizaje cooperativo sobre el individual en el rendimiento académico de estudiantes universitarios, pero no precisan la magnitud obtenida en este subgrupo, investigaciones todas que se analizarán en las próximas páginas y deberán ser discutidas con precaución a la luz de los resultados obtenidos.

Por lo tanto, el objetivo del presente meta-análisis ha sido valorar estadísticamente la eficacia del aprendizaje cooperativo en comparación con el aprendizaje individual en el rendimiento académico de estudiantes universitarios y de ser eficiente, bajo qué condiciones se podría implementar esta metodología de enseñanza en aulas universitarias.

De un total de 3744 estudios registrados en fuentes primarias (2960) y secundarias (784) bajo los descriptores *Cooperative Learning* y *Collaborative Learning*, se han encontrado 2536, de los cuales 2446 se han excluido y solo **90 se han seleccionado según los criterios de inclusión con 119 resultados independientes**, en los que han participado 8114 estudiantes universitarios.

Los 2960 documentos de fuentes primarias aparecen registrados en las bases de datos *Web of Knowledge* (1382), *Periodicals Archive Online* (70), *Proquest dissertation Publishing* (40), en motores de búsqueda como *Google Scholar* y *Google Búsqueda*

*Avanzada* con 389 documentos en formato pdf procedentes de las diez primeras páginas y en revistas especializadas como *European Journal of Social Psychology* (12), *College Teaching* (229), *Higher Education Research Development* (26), *Journal of Educational Psychology* (726), *Research in Higher Educational Journal* (72) y *Journal of College Science Teaching* (14). De estos solo 53 cumplieron con los criterios de inclusión, es decir, un 1,79%.

De los 784 documentos restantes provenientes de fuentes secundarias en meta-análisis (583), otros estudios (165) y revisiones narrativas (36) se han seleccionado solamente 37, lo que equivale a un 4,71%.

### ***Magnitud global del aprendizaje cooperativo***

La magnitud del efecto global sobre el rendimiento académico en estudiantes universitarios, ha sido estadísticamente significativa y positiva a favor del aprendizaje cooperativo frente al aprendizaje individual, con un tamaño moderado en el **modelo de efectos fijos** ( $d=0,56$ ), una  $Q$  estadísticamente significativa con  $Q=430,44$ ,  $p<0.00001$  y una heterogeneidad del 73% y, en el **modelo de efectos aleatorios** ( $d=0,57$ ) con una  $Q=111,2$ ;  $p<0,00001$  y heterogeneidad del 0%.

En ambos modelos se confirma la eficacia moderada del aprendizaje cooperativo en el rendimiento de los estudiantes en comparación con el aprendizaje individual, siendo el modelo de efectos fijos el indicado debido al porcentaje tan elevado de heterogeneidad.

Estos resultados son bastante robustos y así lo corrobora el análisis de sensibilidad porque las magnitudes del efecto al eliminarse una cada vez se mueven entre 0,55 y 0,57, lo que supone un 2% por debajo y por encima respecto a la magnitud del efecto moderado global de 0,56 (modelos de efectos fijos).

De los 90 documentos incluidos se han obtenido 119 resultados independientes (*outcomes*) en donde la magnitud del efecto ha sido positiva y significativa en el 91,59% de todos los estudios individuales, las cuales oscilan entre 0,009 y 2,50 y en solo 10 *outcomes* (8,40%), los efectos han sido negativos a favor del aprendizaje individual pero no estadísticamente significativos.

Sin embargo, es el propio Cohen (1988) quien advierte del peligro de interpretación fuera de contexto de los términos pequeño, mediano y grande, alegación a la que Glass, McGaw y Smith (1981) apoyan cuando argumentan que la efectividad o no de una intervención solo puede afirmarse cuando se compara con otras que buscan producir el mismo efecto.

Parece ser que una característica de las intervenciones educativas es que muy pocos [pocas] tienen efectos que podrían ser descritos por la clasificación de Cohen como *pequeña*. Parece así por los efectos en el rendimiento de los estudiantes. No hay duda que esto es parcialmente un resultado de la amplia variación encontrada en la población como conjunto, en que la estimación de la magnitud del efecto se ha calculado. También uno podría especular que el rendimiento es más difícil de influenciar que otros resultados, quizás porque la mayoría de las escuelas ya están utilizando estrategias óptimas o porque las diferentes estrategias son efectivas en diferentes situaciones (Coe y Merino, 2003, p. 155).

Por tanto, una magnitud del efecto de 0,56 es un resultado consistente con otros meta-análisis que se han centrado solo en el impacto global de esta metodología de enseñanza en la universidad como es el caso de Springer, Stanne y Donovan (1999) con una  $d=0,51$  y el estudio de Johnson y Johnson (2002a) con una  $d=0,53$ . Pero, también este efecto positivo se mantiene, aunque no siempre la magnitud obtenida ha sido moderada, con respecto a otras síntesis cuantitativas en donde en ocasiones la universidad ha sido una variable moderadora mas y no el objeto de estudio.

Tal es el caso de un estudio previo de los hermanos Johnson y Johnson en el año 1989 con estudiantes de todas las etapas educativas, con una magnitud global de 0,64 a favor del aprendizaje cooperativo frente al individual. Cuando se estudian solamente los 137 estudios experimentales de los 226 pertenecientes a la universidad y a adultos, este tamaño continua siendo moderado ( $d=0,62$ ).

O, la pesquisa de Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) al comparar situaciones con y sin agrupamiento en primaria, secundaria y universidad – con solo 7 de un total de 51 investigaciones correspondientes a este último nivel educativo– en donde el efecto global ha sido pequeño con una  $d=0,17$ , resultado no muy diferente al encontrado por estos mismos investigadores en un posterior estudio en donde analizaron los distintos tipos de agrupamiento frente a la instrucción de toda la clase –desde primaria hasta la universidad– con un efecto de 0,16 (Lou, Abrami & Spence, 2000).

Los hallazgos de la investigación de Johnson, Maruyama, Johnson, Nelson y Skon (1981) que tampoco hacen distinción entre primaria, secundaria, universidad y adultos desvelan un efecto global grande ( $d=0,78$ ) mientras que el estudio clásico de Stanne, Johnson y Johnson (1999) en donde se revisaron 754 estudios entre 1900 y 2009 –de los cuales el 41,4% son exclusivos de la universidad y con adultos–, también confirma un efecto positivo pero moderado ( $d=0,64$ ).

Un último ejemplo que vale la pena mencionar pero con población no universitaria es el estudio reciente de Roseth y los hermanos Johnson y Johnson (2008) con adolescentes entre 12 y 15 años de edad quienes revisaron 148 estudios independientes en donde compararon la eficacia de estructuras cooperativas, competitivas e individualistas en el rendimiento y relaciones positivas entre compañeros.

Como predice la teoría de interdependencia social, los resultados indicaron que los mayores logros y relaciones positivas con los compañeros se asociaron con la cooperación en lugar de la competencia o la individualidad. Así, el rendimiento obtuvo una magnitud del efecto mas grande pero moderada entre el aprendizaje cooperativo y el individual ( $d=0,55$ ), seguida por una magnitud pequeña cuando se contrasta la cooperación con la competitividad ( $d=0,46$ ) y entre competitividad e individualidad ( $d=0,20$ ).

### ***Heterogeneidad y variables moderadoras***

Según Higgins (2008) la heterogeneidad entre los estudios es de esperarse y es aceptable si los criterios de elegibilidad que han sido predefinidos son sólidos y los datos son correctos. Dentro de este meta-análisis, se han seleccionado los estudios primarios sobre la base de un conjunto de criterios de inclusión con el fin de aumentar la probabilidad de que las investigaciones sean lo suficientemente similares para obtener una estimación combinada significativa.

En este sentido, no ha sido fácil comparar los resultados de una manera “fiable” y en ocasiones, tampoco la interpretación porque cada estudio se ha llevado a cabo en distintas condiciones (diferentes asignaturas y métodos cooperativos, diferentes instrumentos de medida y duración, entre otros) aunque es tarea de un meta-análisis que esos resultados se combinen como si fueran similares (Eysenck, 1994).



Autores como Hedges y Olkin (1985) consideran que para un análisis de homogeneidad es necesario que cada categoría contenga por lo menos diez observaciones. Esta condición no se ha cumplido en algunas variables explicativas debido al bajo número de estudios por subgrupo, prevaleciendo esta condición en las variables ecológicas. No obstante, otros estudios si reportan mas de diez observaciones pero sigue habiendo una pérdida importante con respecto a la totalidad (90). Tal es el caso de variables metodológicas como el sesgo del investigador (51) y variables ecológicas como número de integrantes (68), clasificación de métodos cooperativos (56), duración del programa (65), conformación de los grupos (48) y estructura de recompensa, de tarea estructura e igualdad de oportunidades para la puntuación (17).

Por consiguiente, el método análogo al análisis de la varianza ANOVA para la variables moderadoras, la prueba  $Q$  propuesta por DerSimonian y Laird (1986), ha generado una baja potencia estadística por lo que los resultados obtenidos tienen que ser interpretados con reserva (Hedges & Pigott, 2004) cuando el número de estudios primarios analizados es pequeño (no sobrepasa frecuentemente el valor de 30 estudios). Más aún, el bajo número de estudios en algunos subgrupos podría deberse fácilmente al azar o a otra variable imprevista. Los estudios pequeños tienden a producir magnitudes del efecto mayores e inestables (Levine, Asada & Carpenter, 2009).

Higgins (2008) sugiere la inclusión de variables como una vía para el análisis de estudios heterogéneos. Y así se ha hecho. Se han estudiado variables moderadoras extrínsecas, ecológicas y metodológicas.

Confirmada la eficacia del aprendizaje cooperativo frente al individual en el rendimiento académico de estudiantes universitarios, el análisis de las variables moderadoras ha permitido estudiar las condiciones eficaces para la implementación de esta metodología de enseñanza en aulas universitarias, sin olvidar que es el propio Slavin (1980) quien afirma que el rendimiento se trata de una variable dependiente compleja de definir conceptual y operacionalmente.

Es una categoría muy heterogénea que podría enmascarar diferencias internas entre las variables (Kyndt et al. 2013) aunque idealmente, todos los estudios utilizados en un metanálisis han de ser ejecutados en las mismas circunstancias experimentales.

Como antesala, a todo el análisis que prosigue en las próximas páginas y reconociendo el bajo número de estudios reportados en algunas de ellas, no se pretende generalizar los resultados encontrados porque

When the goal is only a quantitative integration there is no minimum; two studies is enough, although with a small number of studies the statistical analyses are rather unstable. However, when the goal is to summarize and communicate the “state of the art” and only 3 or 4 studies are available, then it is probably better to make a description and assessment of the studies one by one, indicating the differences and similarities (Botella & Gambara, 2006, p. 429)

### VARIABLES EXTRÍNECAS

*Años de publicación.* Los años de publicación de los estudios incluidos abarcan desde 1980 hasta 2012, ambos inclusive, concentrándose el mayor porcentaje en la década de los 90 con un 33,33% cuando el aprendizaje cooperativo está en pleno auge (Johnson & Johnson, 1999c; Slavin 1999c) aunque actualmente solo entre los años 2010 y 2012 la producción es alentadora con un 18,89%.

El estudio de Akdemir y Arslan (2012) confirma este recorrido del aprendizaje cooperativo como metodología de enseñanza en la literatura científica. Analizaron 2037 documentos encontrados bajo el descriptor *Cooperative Learning* en la base de datos ERIC entre los años 1996 y 2010. Es en el año 1988 cuando aparece el primer documento y en la medida en que avanzan los años, el número de publicaciones aumenta alcanzando el mayor porcentaje entre los años 2005 y 2009.

En el presente meta-análisis, los hallazgos indican que entre los años 2000 y 2012 las magnitudes del efecto son mas grandes ( $d=0,73$ ) que cuando se comparan con los estudios publicados entre los años 1980 y 1999 ( $d=0,43$ ) y, si de lustros se trata, los años 2010–2015 –recordemos que los criterios de inclusión indican solo estudios hasta el 2012 inclusive–, es el período de tiempo que reporta el tamaño mas grande con un 0,88 con una prueba de homogeneidad estadísticamente significativa intragrupos ( $Q_w=365,7$ ;  $p<0,0001$ ) e intergrupos ( $Q_b=64,8$ ;  $p<0,0001$ ).

Al respecto, Qin, Johnson y Johnson (1995) si hallaron diferencias significativas en años no tan recientes en cuanto a la década de publicación en estudios con situaciones competitivas e individuales, entre los años 1900 y 1995 ( $F(5, 57)=4,36$ ;  $p<0,002$ ) pero el número reducido de estudios antes de los años 1960 resta relevancia a esta conclusión porque el 80% de estos documentos se concentra entre 1970 y 1995.

*Tipos de fuentes y documentos.* La mayoría de los estudios primarios que se han incluido en el meta-análisis provienen de fuentes primarias (58,89%) como bases de datos, motores de búsqueda y revistas especializadas quedando en un segundo lugar en orden decreciente las revisiones secundarias procedentes de otros meta-análisis (32,22%), de la lectura de las referencias bibliográficas de otros estudios (7,78%) y con solo un documento incluido, una revisión narrativa (1,1%).

Son las revistas científicas (75,56%) seguidas por las disertaciones y tesis doctorales (16,67%) y los contados reportes técnicos (7,78%) los tipos de documentos que prevalecen, proporciones estas que también se mantienen muy similares en la síntesis cuantitativa de Johnson y Johnson (2002a) para quienes el mayor porcentaje recae en los artículos de investigación (81,1%), luego las tesis doctorales (8,7%) y las de máster (1%) y en último lugar los reportes técnicos (5,4%).

A raíz del meta-análisis efectuado, los reportes técnicos, artículos de investigación, disertaciones y tesis doctorales obtuvieron efectos moderados de 0,58, 0,57 y 0,56, respectivamente. La similitud en los tamaños de los tres subgrupos difiere de Springer, Stanne y Donovan (1999) quienes si hallaron una puntuación mayor con una magnitud moderada en las revistas científicas ( $d=0,56$ ) en comparación con conferencias en congresos con un efecto pequeño ( $d=0,43$ ) aunque estas diferencias no fueron significativas, hallazgo que también comparten Qin, Johnson y Johnson (1995) cuando contrastan situaciones cooperativas con competitivas no encontrando tampoco diferencias estadísticas según los tipos de documentos ( $F(3, 59)=0,35; p<0,79$ ).

Es el sesgo de publicación un problema que se evidencia con mayor fuerza en las Ciencias Sociales (Dickersin, 1994). Los investigadores tienden a publicar aquellos resultados que son positivos o con efectos significativos, dejando a un lado los efectos negativos o nulos y a la vez, son las mismas revistas científicas y las editoriales quienes incitan y promueven la publicación de estos resultados (Dickersin, 1994; Morales, 1993).

*Lugar de procedencia, revistas científicas y grado de colaboración entre autores e instituciones.* Al igual que los hallazgos encontrados en otras síntesis integradoras (Johnson & Johnson, 2002a) el mayor número de estudios proviene de América, específicamente Norteamérica (67,78%) seguido de Europa (18,895), Asia (10%) y en

los tres últimos lugares África (1,11%), Oceanía (1,11%) y un solo caso entre Asia y América (1,11%).

Las revistas en donde se han difundido mas de un documento primario han sido aquellas con temas generales relacionadas con Educación, Psicología y Psicología Educativa y con temáticas específicas en Química, Enfermería y Tecnología: *Journal of Educational Psychology* (7), *Contemporary Educational Psychology* (3), *The Journal of Social Psychology* (3), *Educational Research and Reviews* (3), *Journal of Chemical Education* (2), *Journal of Experimental Education* (2), *Journal of Research on Computing in Education* (2), *Learning and Instruction* (2) y *Nurse Education Today* (2).

Los grupos de investigación que han publicado con mayor frecuencia en los 90 estudios incluidos han sido liderizados por Steven Crooks, Donald Dansereau, John Fantuzzo y Kemal Doymuş dato que es interesante cuando se compara con otras revisiones realizadas solo en la base de datos ERIC y no en la WOK en periodos de tiempo muy similares como la de Akdemir y Arslan (2012) donde ninguno de estos autores aparece entre los 10 primeros.

Sin embargo, cuando se analiza el grado de colaboración entre autores es Roger Johnson quien ha publicado con un gran número de autores en contraposición con Kemal Doymuş quien aunque igualmente es investigador principal su grupo de trabajo es mas “cerrado”. Con respecto a las instituciones, *Password Community Mentoring* en Indianápolis es la que ha colaborado en mayor medida con otras universidades mientras que *Texas Christian University* presenta un elevado número de publicaciones pero con un pequeño grupo de colaboradores.

## VARIABLES METODOLÓGICAS

*Asignación.* La mayoría de los estudios incluidos han empleado la aleatorización de sujetos, grupos o ambos (65,55%) antes que la no aleatorización (34,44%) siendo una condición fundamental de los diseños experimentales pero no así de los cuasi-experimentales para atender así uno de los problemas mas críticos en investigación: el sesgo de selección (Slavin, 1992b).

En aquellos estudios sin asignación aleatoria la magnitud encontrada ha sido mayor con un 0,66 que aquellos con aleatorización 0,50, ambos tamaños muy similares a los resultados de Springer, Stanne y Donovan (1999) con 0,65 y 0,46 respectivamente, pero no así en el estudio de Qin, Johnson y Johnson (1995) quienes no hallaron diferencias

en la asignación al azar entre situaciones competitivas e individuales ( $F(2, 60)=0,98$ ;  $p<0,38$ ).

*Situación basal.* La falta de equivalencia o la no garantía de los grupos antes de la intervención, incluso cuando la aleatorización está presente en diseños experimentales, atenta a las puntuaciones del meta-análisis (Huedo, 2006; Ortego & Botella, 2010) pero, en la presente investigación esta condición mantuvo la misma magnitud del efecto en aquellos estudios, un 80%, con una equivalencia inicial en el grupo control y experimental ( $d=0,59$ ) y también cuando no se avala esta igualdad en un 20% de los documentos antes de comenzar con la intervención ( $d=0,51$ ).

*Sesgo del investigador.* Casi la mitad de los estudios primarios, un 43,33% no ofrecen información relativa al sesgo del investigador y del 56,67% restante, en el 36,67% es la misma persona –por lo general es el propio investigador– quien asume la tarea de aplicar la intervención en el grupo control y experimental.

Es esta una variable de confusión porque “In experiments in which one treatment is compared to another, the critical task is to eliminated by proper research design all sources of differences between the treatment groups other than the effects of the treatments themselves” (Slavin 1992b, p.108) siendo el profesor una variable en si mismo. Sus habilidades u otras características personales influyen conscientemente o no a favor de uno de los grupos de intervención viéndose afectados los resultados del estudio.

Tan es así que los resultados encontrados señalan que cuando es el mismo investigador la persona a cargo del grupo control y experimental, el efecto es moderado ( $d= 0,58$ ) pero esta magnitud disminuye a un tamaño pequeño ( $d=0,48$ ) cuando han sido diferentes investigadores quienes han conducido el estudio, diferencias que también comparte el meta-análisis de Springer, Stanne y Donovan (1999) pero con una disimilitud aun mas acentuada porque varía de una magnitud mas grande ( $d=0,73$ ) a una pequeña ( $d=0,41$ ).

*Instrumentos de medida.* Es conocido que el rendimiento académico puede ser cuantificado con distintos instrumentos de medida y que cualquier intento por sintetizar los resultados de diferentes estudios, debe prestar atención a la validez y fiabilidad de las diversas formas que se utilizan para medir la variable dependiente (Kyndt et al. 2013; Slavin, 1980).

A sabiendas de esta afirmación, el porcentaje de estudios que han o no aplicado instrumentos estandarizados ha sido muy similar. Un 51,64% para los primeros y 45,05% para los segundos.

Con proporciones muy parecidas, la magnitud del efecto en ambos casos es moderada pero es un poco mayor en los estandarizados ( $d=0,62$ ) que en los no estandarizados ( $d=0,55$ ). En cambio, Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) encontraron tamaños pequeños con una  $d=0,34$  en instrumentos elaborados por el investigador y una  $d=0,42$  cuando han sido diseñados por el docente y una  $d$  menor con 0,07 en aquellos estandarizados siendo el instrumento de medida un predictor significativo en el rendimiento ( $Q_b(2)=70,02$ ;  $p<0,5$ ) y, también Springer, Stanne y Donovan (1999) hallaron resultados muy parecidos a estos últimos. Cuando las medidas son exámenes o pruebas no validadas las magnitudes son mayores ( $d=0,59$ ) en comparación con instrumentos estandarizados ( $d=0,33$ ).

*Calidad metodológica.* Algunos autores abogan porque la calidad sea una variable moderadora mas a estudiar (Glass, McGaw & Smith, 1981; Smith, 1980) mientras que otros consideran que debería ser un criterio de inclusión/exclusión de los estudios (Slavin, 1986; Wortman & Bryant, 1985).

En la presente revisión meta-analítica, la calidad metodológica ha sido otra variable de interés. De los 90 estudios primarios, en orden decreciente, el 57,77% ha obtenido una calidad media, un 31,11% reporta una calidad alta y solo el 11,11% una calidad alta.

Si la magnitud global del efecto en estudios sin calidad ha sido moderada con 0,57 (modelo de efectos fijos) este tamaño se conserva igualmente en estudios con media ( $d=0,52$ ) y alta calidad ( $d=0,58$ ) aunque el nivel bajo se mantiene en el mismo rango pero aumenta a 0,74. Resultados similares a estos hallazgos son las conclusiones del estudio de Johnson y Johnson (1989) y Stanne, Johnson y Johnson (1999) en donde el tamaño moderado tampoco varía en estudios con ( $d=0,64$ ) y sin ( $d=0,61$ ) calidad metodológica.

Pero, cuando se comparan los estudios con y sin calidad metodológica en los dos únicos niveles de calidad de los que se dispone información para el cálculo de la magnitud del efecto (nivel medio y alto), los hallazgos indican que no se ven afectados por la calidad los años de publicación entre 1980-1999 y 2000-2012, los métodos cooperativos por parejas *Dyadic Methods* y los que se apoyan con tecnología *Computer Support* y las

áreas de Educación y Ciencias Sociales, Educación Comercial y Derecho mientras que si sufrieron cambios las Ciencias, las Humanidades y Artes, los métodos cooperativos combinados y los métodos *Student Team Learning*.

Solamente en Ciencias con niveles altos de calidad se obtuvieron magnitudes grandes, a lo que alegan Marín, Sánchez–Meca y López (2009) “Cuando los resultados de los estudios cambien en función de su calidad metodológica, se entenderá que los resultados más válidos y fiables serán los procedentes de los estudios con la mayor calidad metodológica” (p.113).

### VARIABLES ECOLÓGICAS

*Asignaturas y áreas de conocimiento.* La implementación del aprendizaje cooperativo en distintas disciplinas evidencia la versatilidad de esta metodología con un total de 15 asignaturas, en donde dominan las clases impartidas sobre Lingüística (14,29%), Matemáticas (12,61%), Química (11,76%), Psicología (9,24%) y Estadística (7,56%).

En las 12 asignaturas estudiadas –por contar estas con al menos mas de un *outcome* para el cálculo de la magnitud– los efectos han sido positivos y a favor del aprendizaje cooperativo, con tamaños que oscilan entre 0,25 y 1,08: Psicología (0,25), Ingeniería (0,34), Economía (0,43), Tecnología (0,46), Administración (0,48), Lingüística (0,53), Matemáticas (0,58), Educación (0,7), Estadística (0,83), Química (0,84), Física (1,02) y Enfermería (1,08).

Si se compara de manera descriptiva solamente los rangos del efecto de los estudios individuales en la asignatura de Tecnología con los hallazgos del estudio de Lou, Abrami y D’Apollonia (2001) por ser la única información detallada de la universidad que ofrecen estos autores, los valores obtenidos son bastante similares. En la presente investigación oscilan entre -0,139 a 1,43 y en el estudio de estos autores es entre -0.04 y valores grandes con 1,40.

Específicamente en Química con 12 estudios analizados y con una magnitud grande del efecto de 0,84, los resultados encontrados difieren significativamente del estudio de Bowen (2000) con 15 investigaciones con un tamaño de 0,37 en estudiantes de bachillerato y universitarios. Como también se distinguen los hallazgos de Kalaian y Kasim (2014) interesados solo en el rendimiento en Estadística en agrupaciones bajo los principios de *Cooperative Learning*, *Collaborative Learning* y *Inquiry–Based Learning* quienes concluyen su investigación con una magnitud moderada de 0,60 (n=10) en

contraposición a un efecto moderado–grande de 0,80 en los 6 documentos primarios encontrados.

Y, si se contrastan los resultados obtenidos solo en Matemáticas con otras síntesis por ejemplo la de Springer, Stanne y Donovan (1999) los valores son similares a los encontrados ( $d=0,58$ ) con un efecto moderado para estos autores de 0,53.

En esta misma línea, otras revisiones cuantitativas fuera del nivel universitario, como la revisión de Stoner (2004) con estudiantes entre 10 y 14 años ha obtenido en Matemáticas una magnitud del efecto pequeña de 0,135 ( $n=23$ ) a favor del aprendizaje cooperativo frente al individual y por su parte, Haas (2005) con alumnos de la ESO también encontró en Álgebra un efecto positivo pero pequeño de 0,34 ( $n=3$ ).

Otro dato interesante a analizar es cuando se agrupan las asignaturas por áreas de conocimiento. En Ciencias –Ciencias de la vida, Ciencias Físicas, Matemáticas, Estadística y Tecnología– los hallazgos encontrados señalan una magnitud moderada de 0,71 a diferencia de los resultados de Springer, Stanne y Donovan (1999) con un efecto pequeño de 0,42 y de Romero (2009) con otro tamaño igualmente pequeño de 0,38 en estudiantes de secundaria y universitarios quien afirma que

No other meta-analyses specifically addressing the impact of cooperative learning on science achievement have been published since Bowen (2000), and he only focused on chemistry achievement. Scott, et al. (2005) reported effects of cooperative learning, but their study focused on other interventions as well and only cited 3 studies on cooperative learning. Therefore, there have been no meta-analyses focusing solely on the impact of cooperative learning on science achievement (including all science disciplines) since the Springer, et al. (1999) study (p. 5).

Cuando se comparan las Ciencias con otras áreas de conocimiento, estas apuntan a una magnitud mayor de 0,71 mientras que por ejemplo en Humanidades y Artes –Música y Lingüística– disminuye a 0,52. En los resultados de Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) esta diferencia a favor de las Matemáticas y Ciencias sobre las Artes también se mantiene pero con efectos mas pequeños (0,20 y 0,13 respectivamente).

En meta-análisis, como el de Schroeder, Scott, Tolson, Huang y Lee (2007) en donde el objeto de estudio es la eficacia del aprendizaje cooperativo en el rendimiento como una de las tantas estrategias docentes que emplean los profesores que imparten Ciencias –Biología, Química, Física, Ciencias de la Tierra y Ciencias en general– a estudiantes desde primaria hasta la ESO ( $n=16$ ) y en bachillerato ( $n=45$ ), ocupa el segundo lugar



con una magnitud grande de 0,96 bajo el nombre de *Collaborative Learning Strategies*, luego de otras como *Enhanced Context Strategies* ( $d=1,48$ ) y antes que *Questioning Strategies* ( $d=0,74$ ), *Inquiry Strategies* ( $d=0,65$ ), *Manipulation Strategies* ( $d=0,57$ ), *Assessment Strategies* ( $d=0,51$ ), *Instructional Technology Strategies* ( $d=0,48$ ) y *Enhanced Material Strategies* ( $d=0,29$ ).

No obstante, los autores aclaran que este resultado tan positivo en la magnitud del efecto en la estrategia objeto de interés – *Collaborative Learning Strategies*– pertenece solo a tres documentos primarios, siendo necesario un total de 55 estudios para disminuir este tamaño del efecto por debajo de 0,05.

*Métodos cooperativos.* El volumen de estudios que no reporta el método cooperativo empleado es llamativo por ser esta una condición nuclear en la comprensión del aprendizaje cooperativo. Solo el 61,34% cumple con las exigencias propias de cada técnica y de estas, sobresalen el trabajo por parejas *Scripted Cooperation* (15,97%), el apoyo de la tecnología *Computer Support* (14,29%) y *Student Teams–Achievement Divisions* (7,56%).

Es el estudio clásico de Johnson, Johnson y Stanne (2000), el único hasta ahora conocido que se ha especializado en la efectividad de varios métodos cooperativos donde se obtuvieron efectos positivos en *Learning Together*, *Group Investigation*, *Team Assited Individualization*, *Student Teams–Achievement Divisions* y *Jigsaw I*. Este dato afirmativo, que también coincide con los resultados encontrados en la presente síntesis cuantitativa difiere cuando se profundiza en cada una de las magnitudes por subgrupo.

De esta forma, los métodos *Student Teams–Achievement Divisions* ( $d=0,32$ ;  $n=6$ ) y *Team Assited Individualization* ( $d=0,36$ ;  $n=2$ ) que han obtenido un efecto pequeño concuerdan con los hallazgos de estos autores para quienes, la magnitud del primero es 0,29 y la del segundo 0,33. En cambio, el método *Jigsaw I* con un efecto moderado ( $d=0,69$ ;  $n=2$ ) y *Group Investigation* con un efecto pequeño–moderado ( $d=0,49$ ;  $n=2$ ) han cambiado en el estudio de Johnson y sus colegas a un efecto pequeño 0,13 y moderado respectivamente 0,62. No obstante, no hay que olvidar que se tratan de hallazgos globales del trabajo de Johnson, Johnson y Stanne (2000) con estudios desde primaria hasta la universidad y que este último nivel solo cuenta con 37 estudios de un total de 158.

*Duración.* El periodo de tiempo para la implementación del aprendizaje cooperativo es muy variado. Los profesores aplican esta metodología en sesiones muy breves de dos o tres días (16,67%) y, durante el semestre de clase, prefieren intervenciones de duración media, entre 5 y 8 semanas (18,89%) y media-alta entre 13 y 16 semanas (7,78%) y en menor proporción con intervenciones entre 9 y 12 semanas (15,56%), entre 1 y 4 semanas (10%) y, las que menos las que sobrepasan las 17 semanas (3,33%) aunque un 27,78% de los estudios no ofrece esta información.

El rendimiento no alcanza su mayor tamaño del efecto con intervenciones muy cortas o puntuales ( $d=0,3$ ) pero si con programas de duración no tan cortos entre 1 y 4 semanas con un efecto grande de 0,79. En periodos de tiempo intermedios, es decir, entre 5 y 8 semanas el efecto es moderado ( $d=0,72$ ) pero a medida que avanza el semestre, entre 9 y 12 semanas y entre 13 y 16 semanas este efecto moderado se mantiene pero decrece (0,5 y 0,51 respectivamente).

Según Springer, Stanne y Donovan (1999) la tendencia genera un mejor rendimiento con efectos grandes ( $d=0,73$ ) en intervenciones medianas –entre 1 y 11 semanas– y un rendimiento moderado en tiempos mas largos, mayores de 11 semanas ( $d=0,52$ ) y en intervenciones puntuales ( $d=0,52$ ), lo que permite observar coincidencias y diferencias entre el meta-análisis realizado y el estudio de los autores citados.

*Conformación de los grupos.* A pesar de ser otra variable crucial para la estructuración y comprensión del funcionamiento de los grupos cooperativos siendo a la vez una condición asociada al método cooperativo que se quiere emplear, el 46,67% de los estudios no reporta esta información. Del 53,34% restante, predominan los grupos heterogéneos con un 35,56% en contraste con un 17,78% de grupos homogéneos.

Si los miembros de los grupos están conformados por grupos homogéneos según niveles de rendimiento similares o características como la edad y el sexo, el efecto es menor con un tamaño pequeño de 0,41 en comparación con grupos heterogéneos con una magnitud moderada de 0,56.

Este resultado se asemeja al estudio de Springer, Stanne y Donovan (1999) con magnitudes muy similares ( $d=0,39$  y  $d=0,55$ , respectivamente) y a la investigación de Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) quienes también encontraron un efecto pequeño pero mucho mas reducido de 0,12 en donde la efectividad de los grupos homogéneos es mayor a los heterogéneos no siendo

significativa esta diferencia ( $Q(19)=43,90$ ;  $p<0,5$ ), aunque no hay que olvidar que los resultados de este análisis son generales porque no ofrecen estimaciones por etapas educativas.

*Número de integrantes.* Los primeros meta-análisis centrados en el tamaño de los grupos focalizaron su atención en el tamaño de las clases y su incidencia en el rendimiento académico de los estudiantes para luego dar paso al estudio de las agrupaciones dentro del aula. Es la revisión cuantitativa de Slavin en 1989b con solo 8 estudios la que sirve de precedente para afirmar que a menor número de estudiantes en primaria en clases reducidas de 20 alumnos, el efecto es positivo en el rendimiento aunque pequeño ( $d=0,13$ ). De esta forma, los meta-análisis posteriores se han dedicado a estudiar el efecto del número de integrantes por grupos cooperativos.

Es el trabajo por parejas el favorito (36,13%), seguido por agrupaciones de cuatro miembros (11,76%), tríos (10,08%), de cinco (6,72%) y de entre cuatro y cinco (6,72%), no siendo estas combinaciones las únicas pero si las que representan el 71,41% del total.

Como las posibilidades de agrupación son muy diversas, solo se han comparado las mas comunes como parejas y tríos que han obtenido una magnitud baja (0,45 y 0,43, respectivamente) siendo los grupos de entre cuatro y seis integrantes los que han alcanzado un efecto moderado de 0,63. El comportamiento de lo que ocurre en grupos conformados por parejas y tríos es similar a otros estudios como el de Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) aunque estos investigadores encontraron efectos todavía mas pequeños con una  $d=0,15$  para dos miembros y entre 3 y 4 estudiantes una  $d=0,22$ .

Mas allá, del estudio del aprendizaje cooperativo como metodología de enseñanza pero bajo los principios de la cooperación y ayuda entre iguales, el estudio de Graham y Perin (2007) midió once distintas estrategias docentes para la enseñanza de la escritura en adolescentes, una de ellas, *Peer Asisstance* en donde los estudiantes trabajan juntos para planificar proyectos y/o revisar sus composiciones. Ocupó el tercer lugar con una magnitud alta de 0,75 siendo la única estrategia de todas en donde la relación es estudiante-estudiante y no, estudiante-profesor, dato que pone de manifiesto los beneficios de la cooperación entre pares.

*Estructura de tareas, recompensas e igualdad de oportunidades para la puntuación.* Si en general, como se ha demostrado, resulta difícil encontrar un número de estudios individuales en ciertas variables moderadoras que permitan obtener la magnitud del efecto, no lo es menos cuando se trata del análisis de los métodos cooperativos según estas estructuras.

Un 81,52% de los estudios no reportan información relativa a la estructura de tareas, recompensas e igualdad de oportunidades para la puntuación porque se tratan de condiciones que se asocian directamente con algunos tipo de métodos cooperativos empleados, a saber *Student Teams–Achievement Divisions*, *Jigsaw I*, *Jigsaw II*, *Learning Together*, *Group Investigation*, *Cooperative Integrated Reading and Composition* y *Teams–Games–Tournament*, aunque de estos dos últimos no se registraron estudios. Además, algunos de los subgrupos que conforman cada estructura tampoco cuentan con un número de estudios mínimos para el cálculo de las estimaciones.

Cuando se comparan los tres tipos de estructura, una sola arroja una diferencia muy marcada: la especialización de la tarea.

Los métodos cooperativos con un alto nivel de especialización en las tareas –*Jigsaw I*, *Jigsaw II*, *Group Investigation*– han obtenido una magnitud grande ( $d=0,88$ ) en comparación con los otros que no exigen tal grado de especialización ( $d=0,36$ ) como son *Student Teams–Achievement Divisions*, *Team Assited Individualization* y *Learning Together*.

En cambio, en cuanto a la estructura de recompensas tanto aquellos que se fundamentan en el esfuerzo individual dentro del grupo *Student Teams–Achievement Divisions*, *Team Assited Individualization* y *Jigsaw II* como los que no requieren de recompensas *Group Investigation* y *Jigsaw II* han obtenido un efecto moderado (0,55 y 0,61, respectivamente).

Lo mismo ocurre con métodos que trabajan bajo la igualdad o no de oportunidades para la puntuación grupal en donde ambos subgrupos también han obtenido una magnitud del efecto moderada. Los primeros, *Student Teams–Achievement Divisions*, *Team Assited Individualization* y *Jigsaw II* con una  $d=0,55$ . Los segundos, *Learning Together*, *Group Investigation* y *Jigsaw II* con una  $d=0,65$ .

Un estudio no tan reciente es el de Slavin (1983) quien en años anteriores se interesó por el análisis de estas variables. De un total de 27 estudios, cuyos métodos no tienen una estructura de tareas especializadas pero la recompensa es individual dentro del grupo (como son *Student Teams–Achievement Divisions*, *Teams–Games–Tournament*, *Team Assisted Individualization*), en 24 de ellos, los efectos son positivos en el rendimiento cuando las recompensas son individuales dentro del grupo, y solo en 3, no se encontraron diferencias ni a favor ni en contra. En cambio, en los 4 estudios en donde los métodos cooperativos sin especialización de la tarea aplican recompensas grupales como el método *Learning Together*, en 3 no se encontraron diferencias y en uno solo el efecto ha sido negativo.

Cuando los métodos cooperativos exigen una especialización de la tarea con recompensas grupales, como es el caso del método *Group Investigation*, de un total de 4 estudios, todos los efectos han sido positivos y solo en uno no se hallaron diferencias. En el caso de métodos con especialización de la tarea pero con recompensas individuales dentro del grupo, el único estudio perteneciente al método *Jigsaw II* reportó un efecto positivo (Slavin, 1983).

En consecuencia, para Slavin (1983) los resultados de esta síntesis cuantitativa apuntan a la eficacia de métodos no especializados en la tarea con una recompensa individual dentro del grupo y métodos especializados con recompensas grupales –resultados estos dos últimos que discrepan de los hallazgos encontrados– y que confirma Gavilán (2004) cuando comenta que

según la investigación desarrollada sobre los sistemas de puntuación en el aprendizaje cooperativo, los que producen mejores resultados son aquellos que dan igualdad de oportunidades a todos los participantes, valorando no solo el rendimiento académico sino también el esfuerzo realizado, de manera que todos los estudiantes dispongan de las mismas oportunidades para contribuir con su actuación a la puntuación del grupo (p.24).

El estudio de Lou, Abrami y D'Apollonia (2001) es otra de esas excepciones en el ámbito universitario que contrasta los resultados antes mencionados. Cuando midieron el efecto de la informática en el aprendizaje cooperativo e individual encontraron que recompensas individuales para puntuaciones grupales lograron un efecto positivo pero pequeño ( $d=0,16$ ,  $n=178$ ,  $IC_{95\%}=+0,12; +0,20$ ) al igual que en aquellos casos en donde la recompensa es grupal ( $d=0,31$ ,  $n=39$ ,  $IC_{95\%}=+0,20; +0,43$ ).

Recientemente, el estudio de Tomcho y Foels del año 2012 especializado en estudiantes de Psicología con el fin de mejorar la enseñanza de profesores universitarios, encontró que los factores mediadores que inciden en el aprendizaje cooperativo son actividades con un alto nivel de interdependencia positiva entre sus miembros, es decir, con un alto grado de especialización de la tarea ( $k=27$ ;  $Z=0,56$ ;  $95\%CI= [+0,39; +0,74]$ ) en comparación con actividades con un nivel bajo de interdependencia ( $k=10$ ,  $Z=0,30$ ;  $IC_{95\%}=[+0,18; +0,42]$ )  $Q(1)=5,79$ ;  $p=0,16$  y cuando los miembros del grupo evalúan a los otros el efecto es menor ( $k=8$ ;  $Z=0,34$ ;  $IC_{95\%}=[+0,17; +0,51]$ ) que cuando no se realiza formalmente esta evaluación ( $k=29$ ;  $Z=0,53$ ;  $IC_{95\%}=[+0,37; +0,70]$ ),  $Q(1)=2.69$ ,  $p=0,10$ , resultado este último no esperado por los autores.

A manera de cierre, la tabla comparativa 117 resume los resultados obtenidos en contraste con las otras revisiones cuantitativas consultadas.

Resulta interesante observar cómo el meta-análisis de Springer, Stanne y Donovan (1999) –que es uno de los dos estudios que se ha llevado a cabo solo con estudiantes universitarios porque el otro es el de Johnson y Johnson (2002a) del cual no se tiene información con respecto a las variables moderadoras–, es el que mayor coincidencia presenta con los hallazgos encontrados. En cambio, en la medida en que las revisiones meta-analíticas incluyen otros niveles educativos o se centran en la cooperación en áreas específicas de conocimiento, las diferencias aumentan por lo que el argumento de Glass, McGaw y Smith (1981) vuelve a confirmarse cuando dicen que “... particular magnitudes of effect will gain meaning by reference to what is typical in similar circumstances” (p.104).

Tabla 117: Contraste entre resultados obtenidos (magnitud del efecto) y otras síntesis cuantitativas

Variables	Resultados obtenidos <i>d</i>	Similares <i>d</i>	Diferentes <i>d</i>
Tipos de documento			
Reportes técnicos	0,58		Springer, Stanne y Donovan (1999) Congresos 0,43
Artículos de investigación	0,57	<b>Springer, Stanne y Donovan (1999) 0,56</b>	
Asignación			
No azar	0,66	<b>Springer, Stanne y Donovan (1999) 0,65</b>	
Azar	0,55		Springer, Stanne y Donovan (1999) 0,46
Sesgo investigador			
Igual investigador	0,58		Springer, Stanne y Donovan (1999) 0,73
Diferente investigador	0,48		Springer, Stanne y Donovan (1999) 0,41
Instrumentos de medida			
Estandarizados	0,62		Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) Primaria, secundaria y universidad 0,07  Springer, Stanne y Donovan (1999) 0,33
No estandarizados	0,55	<b>Springer, Stanne y Donovan (1999) 0,59</b>	Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) Primaria, secundaria y universidad 0,34 (por investigador) 0,42 (por docente)
Calidad metodológica			
Con calidad	Baja= 0,74 Media=0,52 Alta= 0,58		Johnson y Johnson (1989) Todos los niveles educativos Con calidad=0,64  Stanne, Johnson y Johnson (1999) Universidad y adultos Con calidad= 0,64
Sin calidad	0,57		Johnson y Johnson (1989) Todos los niveles educativos Sin calidad=0,61  Stanne, Johnson y Johnson (1999) Universidad y adultos Sin calidad=0,61
Asignaturas			
Tecnología	-0,139 a 1,43	Lou, Abrami y D'Apollonia (2001) informática -0.04 y valores grandes con 1,40	

Química	0,84		Bowen (2000) Bachillerato y universidad 0,37
Estadística	0,80		Kalaian y Kasim (2014) Cooperative Learning, Collaborative Learning y Inquiry– Based Learning 0,60
Matemáticas	0,53	<b>Springer, Stanne y Donovan (1999) 0,58</b>	Stoner (2004) 10-14 años 0,135  Hass (2005) Secundaria Álgebra 0,34
Áreas de conocimiento			
Ciencias	0,71		Springer, Stanne y Donovan (1999) 0,42  Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) Primaria, secundaria y universidad 0,20  Romero (2009) Secundaria y universidad 0,38  Schroeder, Scott, Tolson, Huang y Lee (2007) Primaria, secundaria y bachillerato 0,96
Humanidades y Artes	0,52		Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) Primaria, secundaria y universidad Artes=0,13
Métodos cooperativos			
STAD	0,32	Johnson, Johnson y Stanne (2000) primaria hasta universidad 0,29	
TAI	0,36	Johnson, Johnson y Stanne (2000) primaria hasta universidad 0,33	
Jigsaw I	0,69		Johnson, Johnson y Stanne (2000) Primaria hasta universidad 0,13
GI	0,49		Johnson, Johnson y Stanne (2000) Primaria hasta universidad 0,62
Duración del programa			
Sesiones puntuales	0,3		Springer, Stanne y Donovan (1999) 0,52
1-4 semanas	0,79	<b>Springer, Stanne y Donovan (1999) 1-11 semanas 0,73</b>	
5-8semanas	0,72		
13-16 semanas	0,51	<b>Springer, Stanne y Donovan</b>	



		(1999) 0,52	
Conformación de grupos			
Homogéneos	0,41	Springer, Stanne y Donovan (1999) 0,39	Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) Primaria, secundaria y universidad 0,12
Heterogéneos	0,56	Springer, Stanne y Donovan (1999) 0,55	
Número de integrantes			
Parejas	0,45		Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) Primaria, secundaria y universidad 0,15
Estructura de recompensa			
Recompensa individual	0,55		Lou, Abrami y D'Apollonia (2001) Informática 0,16
Recompensa grupal	0,61		Lou, Abrami y D'Apollonia (2001) Informática 0,31

Fuente: elaboración propia

Finalmente, debido a la naturaleza propia del meta-análisis, las inferencias causales del presente estudio hay que matizarlas, porque aunque el método utiliza una estadística muy sofisticada, no hay que olvidar que esencialmente se trata de un estudio observacional (Cooper & Hedges, 2009; Lipsey & Wilson, 2001) por lo que, las decisiones prácticas que se asuman a partir de la valoración del efecto vienen dadas por sus consecuencias en cuanto a costos y beneficios relativos porque

En educación, si se pudiera demostrar que al hacer un cambio pequeño y de bajo costo se podría elevar el rendimiento académico en una magnitud del efecto como de 0.1, entonces esto podría ser una mejora significativa, particularmente si la mejora es aplicada uniformemente a todos los estudiantes, y aún más si el efecto fuera acumulativo en el tiempo (Coe y Merino, 2003, p.154).

### ***Elementos eficaces para un aprendizaje cooperativo en la universidad***

La Evidencia Basada en Investigación Educativa es la integración de la demostración empírica junto con la formación continua del profesorado lo que permite la toma de decisiones para la impartición de la enseñanza. En la evidencia empírica recae por un lado, el peso de mostrar las condiciones para la implementación de métodos de enseñanza eficaces y por el otro, la sabiduría y capacidad profesional del docente para emplear estas metodologías con su grupo de estudiantes (Moran, 2004) porque

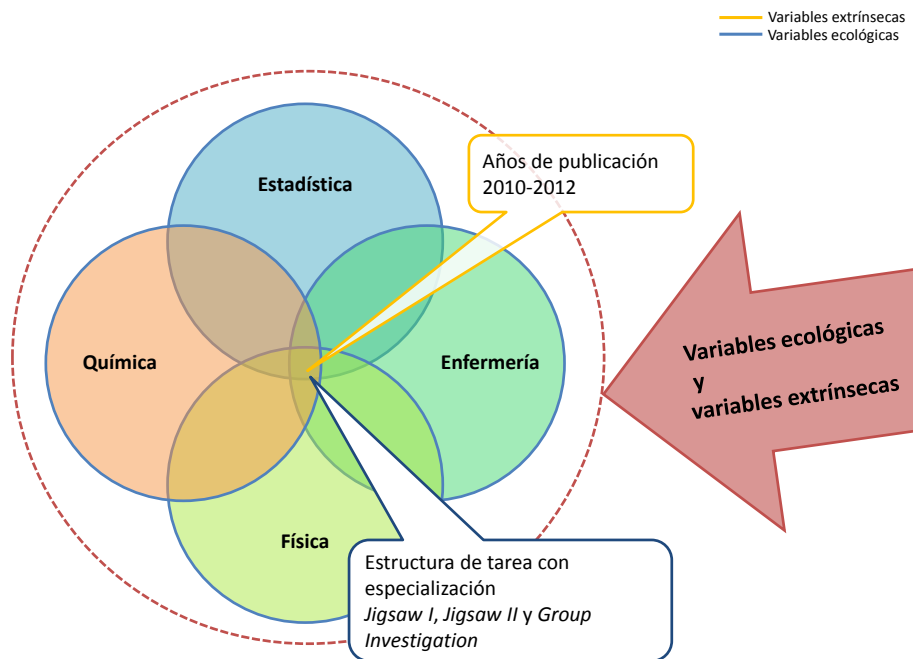
In educational practice, this refers [evidence-based practice in education] to the ability of each University to plan an educational strategy which is supported by the most appropriate teaching and learning research available but interpreted in the light of the needs, values, and current educational level of the students, and the academics' past experiences with similar students (Chipchase, Dalton, Williams & Scutter, 2004, p.133).

De esta forma, los resultados obtenidos en el presente meta-análisis conducen a ofrecer a los profesores universitarios orientaciones para la puesta en práctica del aprendizaje cooperativo bajo condiciones eficaces en el rendimiento académico de los estudiantes.

La figura 16 representa **los mecanismos mas eficientes, es decir, aquellos que han obtenido efectos grandes**, los cuales se concentran exclusivamente en algunos factores del funcionamiento propio de esta metodología (variables ecológicas) y en otras situaciones que aunque escapan al quehacer del profesor en su aula de clase, no dejan de ser menos relevantes para el estudio del objeto de interés (variables extrínsecas) quedándose fuera de esta propuesta las variables relativas al método y al diseño de la investigación (variables metodológicas) en donde no ha prevalecido ninguna con tamaños grandes.

Las clases que mas se benefician son aquellas relativas a **asignaturas de Estadística, Química, Enfermería y Física**, con **los métodos cooperativos Jigsaw I, Jigsaw II y Group Investigation** que requieren de una especialización de la tarea para lograr los objetivos propuestos y, si de **publicaciones** se trata, las mas recientes han obtenido efectos mas grandes específicamente **entre los años 2010 y 2012** (ilustración 16).

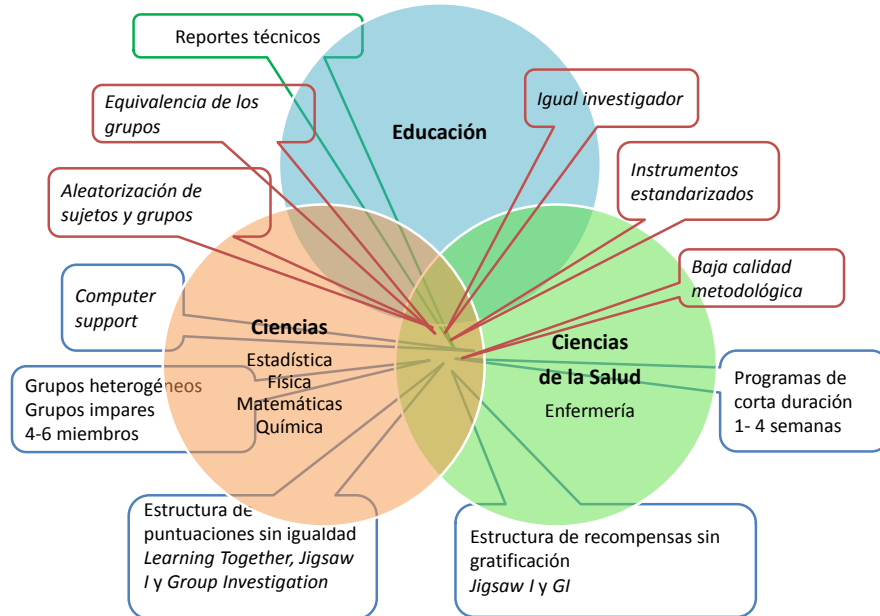
Ilustración 16 Elementos eficaces para un aprendizaje cooperativo en la universidad (efectos grandes)



Fuente: elaboración propia

A estas condiciones óptimas (ilustración 16) se suman **los efectos moderados que prevalecen** en cada una de las variables estudiadas por tratarse de factores que igualmente conciernen al profesor universitario para la toma de decisiones en su aula de clase (ilustración 17).

Ilustración 17: Elementos eficaces para un aprendizaje cooperativo en la universidad (efectos moderados)



Fuente: elaboración propia

**El rendimiento académico de estudiantes universitarios es el mejor** –entendido como el **efecto moderado** que prevalece cuando se comparan distintas variables– **cundo se trabaja el aprendizaje cooperativo** en áreas de conocimiento como **Ciencias** en asignaturas de Estadística, Física, Matemáticas, Química e Informática también en **Ciencias de la Salud** específicamente en Enfermería y en **Educación**, cuando se enseña a docentes en formación Educación Física, Educación en Matemáticas, Educación Musical y Educación para la Lectura y Escritura; con métodos que se apoyan en el uso de la tecnología **Computer Support**, en **programas de baja duración** –entre 1 y 4 semanas–, con **grupos heterogéneos impares conformados entre 4 y 6 miembros** y preferiblemente con aquellos **métodos cooperativos** en donde **no se ofrecen recompensas o gratificaciones al grupo** (*Jigsaw I* y *Group Investigation*) y cuando **no todos los miembros del grupo disponen de las mismas oportunidades para contribuir con su desempeño a la puntuación del grupo** lo que se traduce en calificaciones diferentes para cada integrante (*Learning Together*, *Jigsaw I* y *Group Investigation*) (ver ilustración 17).

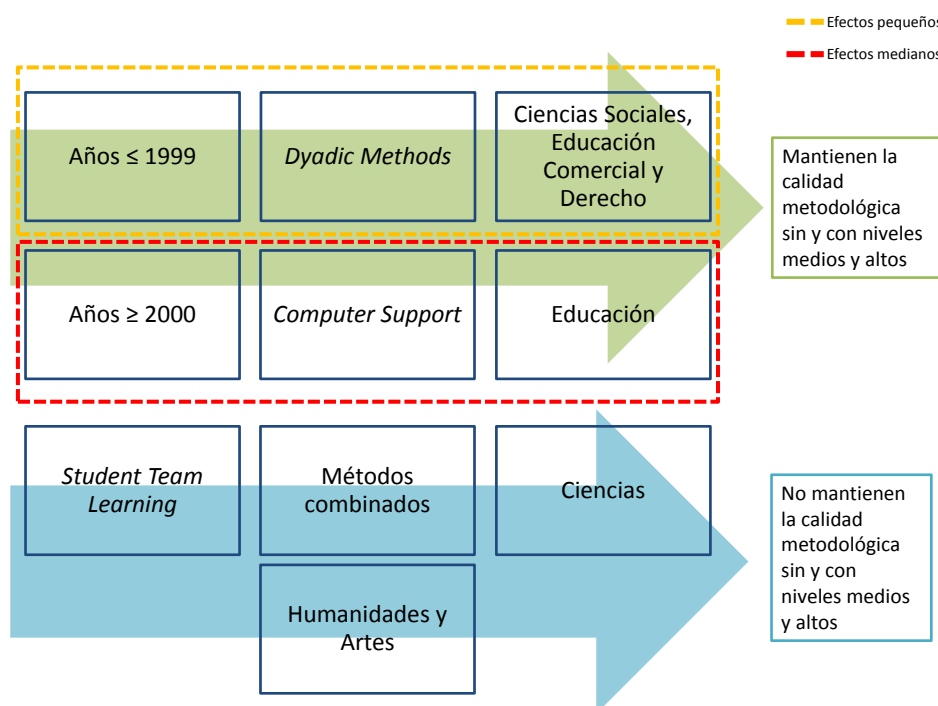
Cuando el profesor universitario decide asumir condiciones metodológicas debe estar al tanto que los efectos moderados que igualmente predominan son el producto de la

**asignación al azar de los sujetos y grupos**; si es **el mismo investigador** el que lleva a cabo la intervención; si se garantiza **la equivalencia de los grupos** antes de la puesta en marcha del programa y si los **instrumentos son estandarizados** y que, los **reportes técnicos** (variable extrínseca) también han sobresalido en las magnitudes del efecto aunque no diferenciándose significativamente de otros tipos de documentos como las disertaciones y artículos publicados en revistas científicas (ver ilustración 17).

Con respecto a la **calidad metodológica** y su incidencia, si los estudios se caracterizan por un nivel bajo de calidad aunque igualmente mantienen un efecto moderado en el rendimiento académico tiende a ser mas elevado en comparación con estudios con una calidad media y alta, por lo que, un análisis de aquellas condiciones que se ven o no afectadas por la calidad de los estudios es otro factor a considerar (ilustración 18).

Por ejemplo, **si se imparten clases en Educación y Ciencias Sociales, Educación Comercial y Derecho o se emplean los métodos *Computer Support* y *Dyadic Methods* se mantienen los efectos en el rendimiento de los estudiantes sin que la calidad sea un factor que incida en los resultados académicos**. Lo mismo ocurre con **documentos publicados entre los años 1980-1999 y entre los años 2000-2012**. Específicamente, si el profesor universitario consulta estudios realizados antes del año 1999 con la finalidad de aplicar las condiciones metodológicas que sus autores sugieren, valorados con una calidad media y alta encontrará tamaños del efecto pequeños en el rendimiento de los estudiantes, pero si revisa investigaciones mas recientes, entre los años 2000 y 2012, asumiendo los mismos criterios de calidad, la magnitud del efecto aumentará a un nivel moderado (ilustración 18).

Ilustración 18: La calidad metodológica y su incidencia en el rendimiento académico en la universidad



Fuente: elaboración propia

En cambio, **hay otras condiciones que sí se ven afectadas por la calidad de los estudios como los métodos *Student Team Learning*, los métodos combinados y las áreas de Ciencias y Humanidades y Artes** (ilustración 18).

La ilustración 19 representa con detalles estos cambios.

Se recomienda con especial atención a los docentes que deseen impartir el aprendizaje cooperativo en asignaturas en **Humanidades y Artes** y cuando emplean **métodos cooperativos combinados** que los efectos serán mas bajos cuando se asume un diseño de investigación de una alta calidad metodológica a diferencia de diseños con calidad media que siguen manteniendo un efecto medio en el rendimiento de los estudiantes universitarios.

Caso contrario a lo que ocurre cuando se imparten métodos que se agrupan bajo el paraguas ***Student Team Learning*** que con una alta calidad se reducen los tamaños a un efecto medio y a su vez, con un nivel de calidad medio se reducen aun mas a tamaños pequeños. Destaca que asignaturas que se imparten en el área de **Ciencias** también varían sus efectos según el nivel de calidad del diseño pero con una marcada diferencia:

con niveles altos de calidad se obtienen magnitudes grandes y con estudios de calidad media se conserva igualmente el mismo nivel (ilustración 19).

Ilustración 19: Niveles de calidad metodológica y magnitudes del efecto

	Magnitud del efecto		
Niveles de calidad	pequeño	moderado	grande
media	STL	Combinados Humanidades y Artes	
alta	Combinados Humanidades y Artes	Ciencias	Ciencias

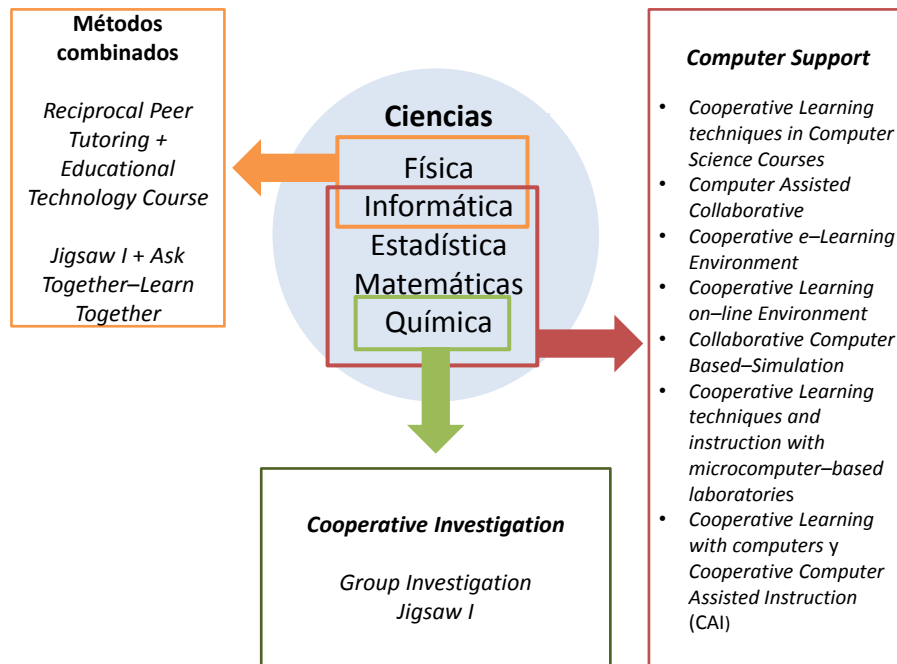
\* no se logró calcular la magnitud del efecto para estudios con un nivel bajo de calidad porque no se contó con suficiente información

Fuente: elaboración propia

Ahora bien, si se precisa por **áreas de conocimiento**, la ilustraciones 20 y 21 recogen el análisis por estratos en Ciencias; Ciencias Sociales, Educación Comercial y Derecho y Humanidades y Artes que son los únicos subgrupos de los que se dispone de información para el cálculo de la magnitud del efecto.

**Las condiciones mas eficientes cuando se aplica el aprendizaje cooperativo en Ciencias –Estadística, Física, Matemáticas, Química e Informática– son aquellas en donde se emplean los métodos combinados para la enseñanza de la Física y la Tecnología (informática) (ilustración 20).**

Ilustración 20: Elementos eficaces del aprendizaje cooperativo para la enseñanza de las Ciencias



Fuente: elaboración propia

Estos métodos combinados potencian el rendimiento al aplicar mas de una técnica cooperativa o al trabajar con otras metodologías de aprendizaje. Ejemplos de estas buenas prácticas son

- (a) **El método por parejas *Reciprocal Peer Tutoring*** cuya dinámica consiste en que cada miembro alterna el rol de tutor y tutelado. El tutor ofrece indicaciones y problemas alternativos para usar frente a su tutelado en caso de que este cometa errores y es **con el apoyo de la tecnología cara a cara o virtual** (*Educational Technology Course*) una posibilidad para incrementar el rendimiento de estudiantes universitarios en formación.
- (b) Otras combinaciones interesantes han resultado de **la composición del *Jigsaw I* con nuevas propuestas de agrupaciones como es la técnica *Ask Together-Learn Together*** que se resume en 10 pasos: organización de los grupos, lectura individual y en silencio del material, elaboración de preguntas individuales relacionadas con la lectura, elaboración de preguntas en grupo, intercambio de preguntas entre los grupos, respuesta grupal a las preguntas que han sido asignadas, presentación de las respuestas por cada grupo a toda la clase,



evaluación de las respuestas grupales, discusión general con toda la clase y por último, examen individual.

A estas recomendaciones se citan otras que tampoco deberían descartarse debido a su eficiencia no tan alta como el caso de los métodos combinados pero igualmente válidas para el área de Ciencias (ver ilustración 20):

(a) **El apoyo de la tecnología *Computer Support* para impartir asignaturas de Tecnología (informática), Química, Estadística y Matemáticas.**

Definitivamente el uso de la tecnología en modalidad presencial o virtual como complemento a otros métodos cooperativos potencia la eficacia. Estos métodos asumen las premisas y principios de la colaboración cara–cara y las adaptan a la tecnología ofreciendo también resultados en el rendimiento académico y en las habilidades sociales.

El aprendizaje colaborativo al ser menos estructurado que el aprendizaje cooperativo adopta variedad de combinaciones y aplicaciones, entre las que se han empleado *Cooperative Learning techniques in Computer Science Courses*, *Computer Assisted Collaborative*, *Cooperative e–Learning Environment*, *Cooperative Learning on–line Environment*, *Collaborative Computer Based–Simulation*, *Cooperative Learning techniques and instruction with microcomputer–based laboratories*, *Cooperative Learning with computers* y *Cooperative Computer Assisted Instruction (CAI)*.

(b) **Los métodos que se agrupan bajo la clasificación *Cooperative Investigation* que incluye a *Group Investigation* y *Jigsaw I* para la enseñanza de la Química.** Se tratan de métodos especializados tanto en contenido y destrezas en donde la evaluación grupal proviene de un proyecto grupal. La dinámica grupal se centra en qué estudiar y cómo se aprende.

La estructura del *Group Investigation* se sustenta en cuatro pilares: investigación, interacción, interpretación y motivación intrínseca. Todo gira en torno a la investigación que tienen desarrollar los estudiantes (subtópicos), la cual se presenta a toda la clase y en donde tanto alumnos como profesores evalúan los proyectos. Por su parte, el *Jigsaw I* exige que los miembros del grupo comprendan que el otro es una fuente de información, en donde la competencia individual es incompatible con el éxito y que, lograr la meta se

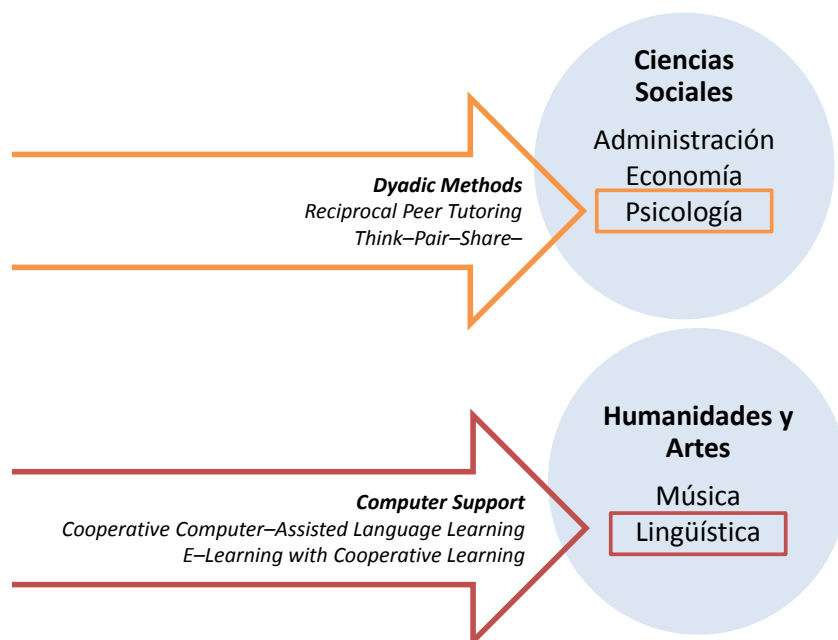
alcanza a través de conductas cooperativas en donde cada miembro aporta al colectivo de compañeros.

Continuando con el análisis por áreas de conocimiento (ilustración 21) pero ahora **en Ciencias Sociales** –Administración, Economía y Psicología– **los métodos cooperativos que son mas eficientes son aquellos que se agrupan bajo la denominación *Dyadic Methods* para la asignatura de Psicología**. Como son solo dos los miembros que conforman el grupo, la responsabilidad del aprendizaje recae con mayor peso en cada estudiante. Las técnicas cooperativas que se recomiendan para el profesor universitario que imparte clases en esta asignatura se resumen en

- (a) ***Reciprocal Peer Tutoring*** porque cada miembro elabora diferentes pruebas sobre el tema, las cuales tienen que ser respondidas por su par, ofreciendo seguimiento en la explicación y resolución de las preguntas antes de la presentación del examen.
- (b) ***Think–Pair–Share***– que se diferencia de las demás técnicas en parejas, por tratarse de un método mucho mas sencillo en donde el docente presenta el tema a la clase, los estudiantes se sientan en parejas y luego acuerdan una respuesta consensuada a partir de las preguntas planteadas por el profesor.

Por último, **en Humanidades y Artes** –Lingüística y Música– **los métodos cooperativos que son mas eficientes son aquellos que se apoyan con la tecnología *Computer Support* en modalidad presencial o virtual para las asignaturas relacionadas con la Lingüística** (ilustración 21). Ejemplos de estas prácticas se evidencian en la aplicación del método *Cooperative Computer–Assisted Language Learning* (CALL) empleado a distancia para el aprendizaje de una lengua extranjera o la modalidad cara a cara cuando se utiliza *E–Learning with Cooperative Learning* (ELCL) para la competencia oral de una segunda lengua.

Ilustración 21: Elementos eficaces del aprendizaje cooperativo para la enseñanza de las Ciencias Sociales y Humanidades y Artes

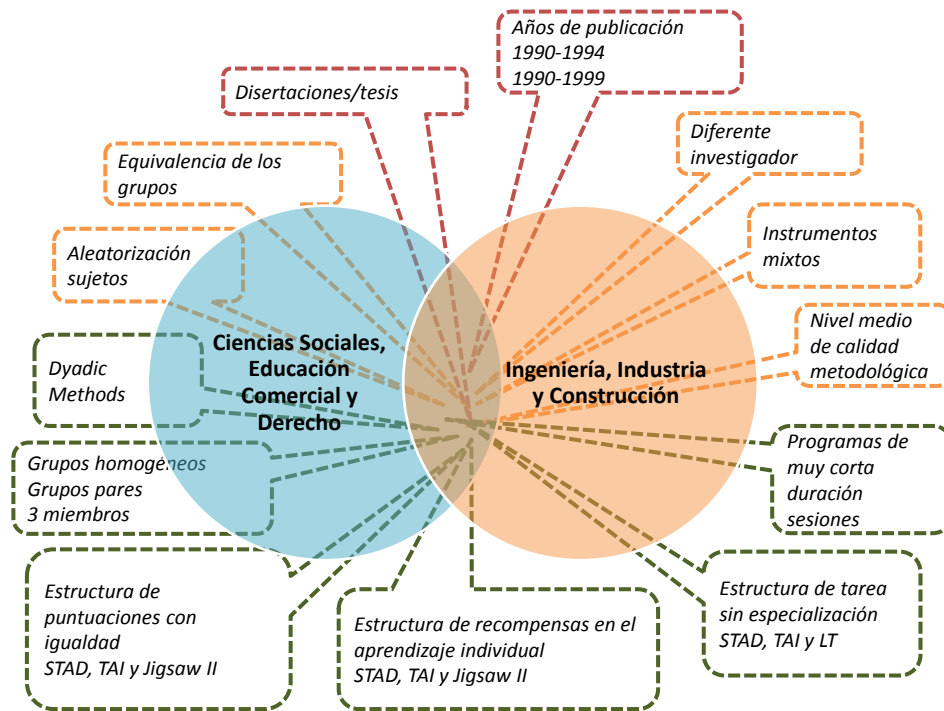


Fuente: elaboración propia

Y lo contrario.

En donde el aprendizaje cooperativo ha sido menos eficaz en la mejora del rendimiento en estudiantes universitarios (ilustración 22), sin que ello no signifique que asimismo ha habido una ganancia cuando se compara con el aprendizaje individual, es cuando se trabaja en Ciencias Sociales, Educación Comercial y Derecho y en Ingeniería, Industria y Construcción; el método *Dyadic Methods*; sesiones cortas de dos o tres intervenciones; grupos homogéneos pares o de tres miembros que son asignados al azar; sin equivalencia inicial entre los grupos; cuando el investigador es diferente; con instrumentos de medición mixtos y por último; con métodos cooperativos que no exigen a cada uno de los miembros una especialización de la tarea (*Student Teams–Achievement Divisions*, *Team Assited Individualization* y *Learning Together*); con una estructura de recompensas basada en aprendizajes individuales (*Student Teams–Achievement Divisions*, *Team Assited Individualization* y *Jigsaw II*) y con igualdad de oportunidades para la puntuación (*Student Teams–Achievement Divisions*, *Team Assited Individualization* y *Jigsaw II*). Son los estudios publicados en el lustro 1990-1994 y en la década 1990-1999 así como un nivel de calidad metodológica media los que reportan los efectos mas pequeños.

Ilustración 22: Elementos menos eficaces para un aprendizaje cooperativo en la universidad



Fuente: elaboración propia

Los hallazgos encontrados en la presente revisión meta-analítica son un pequeño aporte para mejorar la práctica y calidad universitaria porque cuando de innovación educativa se trata, las metodologías de enseñanza son el eje fundamental para el cambio, un cambio que se debe sustentar sobre la base de la evidencia

One of the most important questions in educational research is, “What are the major influences on student achievement?” This question has a voluminous research literature. It would hardly be an exaggeration to say that every hour an article is produced on this topic! The downside of this abundance is that one can find articles to defend almost any position. One of the major problems in education is that “everything seems to work” – we have millions of teachers believing that what they do makes a difference, we have claims galore about the latest thing to make a difference, we have large numbers of articles starting with the common refrain that “evidence shows,” and we have “What works” recipes aplenty. It is not difficult to find studies on this topic that can be used to defend any claim, product, or policy aimed at enhancing achievement (Hattie, Rogers & Swaminathan, 2014, p. 197).



## **CAPÍTULO VI. CONCLUSIONES E IMPLICACIONES EDUCATIVAS**

Una pregunta recurrente en profesores y estudiantes es ¿Por qué cambiar las metodologías de enseñanza–aprendizaje? ¿Por qué aprender de otra manera? En consecuencia ¿Por qué innovar?

Las respuestas no son fáciles a estas interrogantes, pero cuando se toma la decisión de innovar es porque los cambios que se han asumido se sustentan sobre la base de la evidencia, las transformaciones que se contraen están al servicio de la práctica educativa y mejoran tanto el aprendizaje de los alumnos como la eficacia del profesor en su enseñanza.

Los cambios cualitativos que están sufriendo las universidades europeas en el proceso de convergencia en el Espacio Europeo de Educación Superior exigen al profesorado la innovación en metodologías de aprendizaje. Las reformas curriculares y metodológicas apuntan al alumno como centro del proceso de aprendizaje. El estudiante debe ser capaz de responder a los retos que le exhorta la sociedad como ciudadano y futuro profesional y, es el profesor, un incansable pensador sobre su acción apoyada en una permanente reflexión didáctica.

El Plan Bolonia reclama una formación universitaria con metodologías participativas y flexibles, todo un reto para la universidad española que se ha caracterizado por ser mas teórica y menos práctica siendo la etapa educativa donde tradicionalmente ha habido una menor asignación de recursos y donde lamentablemente converge una de las tasas de abandono mayor de Europa.

Es la formación del profesorado la piedra angular de este desafío.

Una de las competencias que se espera del profesor universitario para el ejercicio exitoso de su labor docente es la gestión de métodos de enseñanza o de actividades en conjunto con la planificación del diseño instructivo, el intercambio de información entre el profesor y el alumno durante el proceso de aprendizaje, el seguimiento y la evaluación pero también ofrecer espacios para el trabajo autónomo y en especial para la formación de grupos de trabajo. Igualmente, es el trabajo en grupo una competencia transversal demandada al alumno por estar presente en la mayoría de las titulaciones

universitarias. Es asimismo una competencia profesional que solicita el mercado laboral para la búsqueda de empleabilidad.

El camino para la toma de decisiones en materia de política educativa universitaria es la apuesta por la renovación de metodologías de enseñanza–aprendizaje sustentada en una *Investigación en Educación Basada en la Evidencia*. La idea de que la educación debe ser o convertirse en una práctica basada en la evidencia y que la enseñanza debe ser o debe convertirse en una profesión basada en la evidencia es tema de debate actual. Reducir la brecha entre investigación, política y práctica para una educación de calidad, es el fin.

El aprendizaje cooperativo es un ejemplo de esa innovación metodológica porque las conclusiones provenientes de estudios meta–analíticos desde finales del siglo XX así lo confirman. Se trata de una metodología que no solo favorece el rendimiento académico sino además situaciones afectivas e interpersonales. La versatilidad de sus métodos es un hecho comprobado porque pueden aplicarse en diferentes edades y en una variedad de situaciones y contextos diversos.

Sin embargo, la universidad es una de las etapas educativas que menor número de síntesis cuantitativas reporta, debido a que son escasos los estudios primarios que miden la eficacia de esta metodología porque prevalece la clase magistral, se requiere de formación del profesorado –pero también del estudiante– y de políticas educativas que decanten recursos humanos, físicos y económicos para su puesta en práctica.

Los hallazgos obtenidos en el presente meta–análisis son un pequeño aporte a las contadas síntesis cuantitativas que se han realizado en el ámbito universitario, la última ya hace 12 años, en el 2002a con la finalidad de actualizar las evidencias empíricas en esta área del saber.

El objetivo general del presente estudio ha sido comprobar si estadísticamente el aprendizaje cooperativo es eficaz como metodología de aprendizaje cuando se compara con el aprendizaje individual en el rendimiento académico de estudiantes universitarios.

Los resultados confirman la superioridad del aprendizaje cooperativo sobre el individual en el rendimiento de estudiantes universitarios, con una magnitud del efecto moderada de 0,56 según el modelo de efectos fijos y con una  $Q$  estadísticamente significativa  $Q=430,44$ ;  $p<0,00001$  pero con una heterogeneidad elevada del 73%.

Para el logro de este objetivo general, se ha cumplido con cada uno de los tres objetivos específicos propuestos.

El primer objetivo específico *Integrar cuantitativamente los resultados del aprendizaje cooperativo y del aprendizaje individual de estudios experimentales y cuasiexperimentales que midan el rendimiento académico de estudiantes universitarios* se ha logrado a partir de la revisión de la literatura, el establecimiento de los criterios de inclusión/exclusión de los estudios primarios, su codificación y posterior análisis cuantitativo bajo una métrica común, llamado en español la magnitud global del efecto –en inglés *effects size*–.

Para la revisión de la literatura se han empleado estrategias formales a través de fuentes primarias –bases de datos, motores de búsqueda de Internet y revistas especializadas– y secundarias –otros meta-análisis, estudios primarios y revisión narrativa– no habiendo sido posible complementar la búsqueda de estudios con estrategias informales debido a la no respuesta de autores y grupos de investigación vinculados al área de interés.

Los criterios de inclusión que se han utilizado se resumen en estudios escritos en inglés o en español, entre los años 1980 y 2012, con estudiantes universitarios, de cualquier carrera que cursan estudios superiores, en donde se ha aplicado el aprendizaje cooperativo y el aprendizaje individual, para medir el rendimiento académico, con medidas antes y después o solo después de la intervención y con suficiente información para el cálculo de la magnitud del efecto.

La muestra de estudios que se ha analizado se limita únicamente a diseños experimentales y cuasiexperimentales, no siendo este tipo de diseño el único pero sí el mas apropiado para medir el efecto del tratamiento.

Una ventaja de la utilización de un diseño cuasi-experimental es su fuerte validez ecológica que a la vez podría reducir la validez interna. De los 90 estudios incluidos, en 80 (88,88%) los estudiantes fueron investigados en su entorno natural, lo que contribuye a la generalización de las conclusiones a poblaciones similares.

Como consecuencia, se ha intentado salvar la brecha entre la investigación y la práctica educativa y por ello, los resultados deben interpretarse desde esta perspectiva sabiendo que aunque no se ha establecido como criterio de inclusión que la intervención se debería haber llevado a cabo exclusivamente en aulas ordinarias solo 10 estudios (11,11%) se han diseñado en aulas especiales.



De un total de 3744 estudios registrados en las fuentes primarias y secundarias seleccionadas, bajo los descriptores *Cooperative Learning* y *Collaborative Learning*, se han encontrado 2536, de los cuales 2446 no han cumplido con los criterios de inclusión y solo 90 han sido elegidos con 119 resultados independientes en donde han participado 8114 estudiantes universitarios.

Cada uno de esos estudios primarios ha sido registrado y codificado según sus características en variables extrínsecas, metodológicas y ecológicas por ser consideradas como posibles variables explicativas del efecto. Este proceso de categorización se ha conseguido a través de un manual de codificación.

Una sinopsis de esa codificación se resume en que el período de tiempo que abarcan los años de publicación de los estudios incluidos se enmarcan entre 1980–2012, con el mayor porcentaje de difusión en la década de los 90, la mayoría de fuentes primarias, específicamente de revistas científicas y publicados en Norteamérica siendo Estados Unidos el país de preferencia.

En las decisiones metodológicas asumidas por los investigadores para el diseño de sus estudios ha prevalecido la aleatorización, con equivalencia inicial de los grupos, instrumentos estandarizados, con una calidad media y aunque ha habido una falta de información relevante con relación al sesgo del investigador se impone el mismo investigador para implementar la intervención.

A ese compendio se suma que las asignaturas en que los profesores han impartido con mayor frecuencia el aprendizaje cooperativo han sido Lingüística, Matemáticas, Química, Psicología y Estadística, los métodos cooperativos mas empleados *Scripted Cooperation*, con apoyo de la tecnología *Computer Support* y STAD, con intervenciones de mediana–alta duración (9 y 12 semanas), en donde casi la mitad de los estudios no reporta información de la conformación de los grupos y en los que sí, abundan los grupos heterogéneos, siendo el trabajo por parejas el favorito y con solo el 18,48% de información relativa a la estructura de recompensas, tareas e igualdad de oportunidades para la puntuación.

Y, una vez categorizados los 119 resultados independientes de los 90 estudios incluidos se han logrado sintetizar esos resultados individuales que provienen de distintas investigaciones a través de la estimación global del efecto, que como se ha mencionado anteriormente ha sido positiva, moderada y a favor del aprendizaje cooperativo.

El segundo objetivo específico *Analizar las variables moderadoras del aprendizaje cooperativo que favorecen el rendimiento académico de los estudiantes universitarios* se ha alcanzado a través de un análisis por subgrupos y estratos con resultados que indican efectos positivos en todas las variables moderadoras a excepción de cuando se emplea la asignación de sujetos por emparejamiento con un efecto negativo pero no significativo y en otros dos subgrupos: el lustro 2000–2004 e instrumentos de medidas mixtos ambos con tamaños pequeños del efecto y tampoco significativos.

Sin embargo, cuando se ha analizado con detalle las variables moderadoras, las ecológicas ocupan el primer lugar con resultados contundentes que evidencian las condiciones que mejoran el rendimiento académico de los estudiantes en situaciones cooperativas, en un segundo lugar las variables extrínsecas y en el último lugar, con mucho menos fuerza las variables metodológicas, cuyos efectos en cada uno de los subgrupos siguen siendo positivos pero muy similares no prevaleciendo algunos de ellos.

Estadística, Química, Enfermería y Física son las asignaturas que mas se benefician y los métodos que se sustentan en una estructura especializada de la tarea como son *Jigsaw I*, *Jigsaw II* y *Group Investigation*.

Otro dato relacionado con las variables extrínsecas de los documentos ha sido que las mayores magnitudes del efecto se han encontrado en publicaciones mas recientes entre los años 2000 y 2012, pero que, si se trata de reportes técnicos, tesis doctorales o artículos de investigación, el efecto moderado ha sido similar en los tres subgrupos.

En cuanto a las variables metodológicas, todas las magnitudes del efecto han sido moderadas por lo que no se han encontrado resultados notorios en los subgrupos aunque en este rango, las magnitudes mayores se han hallado en la asignación al azar de sujetos y grupos; cuando se garantiza antes de la intervención la equivalencia entre los grupos, si es el mismo investigador quien lleva a cabo la intervención y si los instrumentos de medida son estandarizados. No obstante, no es de olvidar que estos resultados deben asumirse con precaución porque aunque su incidencia es positiva no prevalecen sobre los otros subgrupos que forman parte de este conjunto de variables.

La calidad de los estudios ha sido otra variable moderadora cuyo tamaño del efecto no ha variado en función de sus tres nivel de calidad. Esta se ha medido a partir de una escala creada *ad-hoc* como resultado de la valoración del juicio de 35 expertos –19

especialistas del ámbito nacional y 16 procedentes del ámbito internacional– con un Alfa de Cronbach de 0,941, versión que siguió siendo mejorada en la redacción de algunos ítems a partir de la fiabilidad inter–observadores, luego de su aplicación en 15 estudios seleccionados al azar.

Cuando se analiza en particular la calidad según los años de publicación, las áreas de conocimiento y los métodos cooperativos, los hallazgos encontrados deben asumirse igualmente con cautela debido al bajo número de estudios registrados en algunos de estos subgrupos.

En los niveles de calidad media y alta mantienen la misma magnitud del efecto que estudios sin calidad los años de publicación entre 1980-1999 y 2000-2012, los métodos cooperativos por parejas *Dyadic Methods* y los que se apoyan con tecnología *Computer Support* y las áreas de Educación y Ciencias Sociales, Educación Comercial y Derecho.

Entre los años 1980 y 1999, tamaños pequeños se mantienen con calidades de nivel medio y alto y efectos medianos con calidades medias y altas entre los años 2000 y 2012.

En los métodos cooperativos que se agrupan bajo la clasificación de *Student Team Learning*, se conservan efectos pequeños en estudios de media calidad pero cambia a un tamaño moderado con estudios con calidades de nivel alto. Igualmente se modifican los métodos combinados con un tamaño moderado en niveles medios de calidad y pequeño cuando es alta.

Y, en Humanidades y Artes las magnitudes también varían. Con niveles altos de calidad el tamaño es pequeño mientras que con una calidad media el tamaño se reduce a moderado.

El tercer y último objetivo específico *Ofrecer orientaciones metodológicas y elementos eficaces para la puesta en práctica del aprendizaje cooperativo* se ha alcanzado precisamente al estudiar en qué condiciones se potencia la eficacia del aprendizaje cooperativo.

Las áreas de conocimiento y asignaturas donde resulta mas eficaz esta metodología han sido Ciencias –Estadística, Química y Física– y Ciencias de la Salud –Enfermería– y con métodos cooperativos que requieren de una especialización de la tarea *Jigsaw I*, *Jigsaw II* y *Group Investigation*.

Específicamente en Ciencias los métodos cooperativos mas eficientes son los que se apoyan con tecnología *Computer Support*, *Cooperative Investigation* y con métodos cooperativos combinados, en Ciencias Sociales los métodos en parejas *Dyadic Methods* y en Humanidades y Artes resulta mas eficaz cuando la cooperación se apoya con la tecnología *Computer Support*.

En consecuencia, la universidad tiene en el aprendizaje cooperativo un mecanismo de enseñanza del trabajo en equipo. Las aulas se convierten en laboratorios de acción y reacción sobre lo existente y en plataformas de estudio sobre lo futuro. Profesores y alumnos pueden, gracias al aprendizaje cooperativo, experimentar, investigar, profundizar, preguntarse y hasta aprender de los errores que se puedan cometer tratando de implementarlo.

Aprender para cooperar, cooperar para aprender. Sin duda lo más excitante del trabajo en grupo está en su habilidad para comprender que las metas alcanzadas son el producto del aporte individual para una producción colectiva.

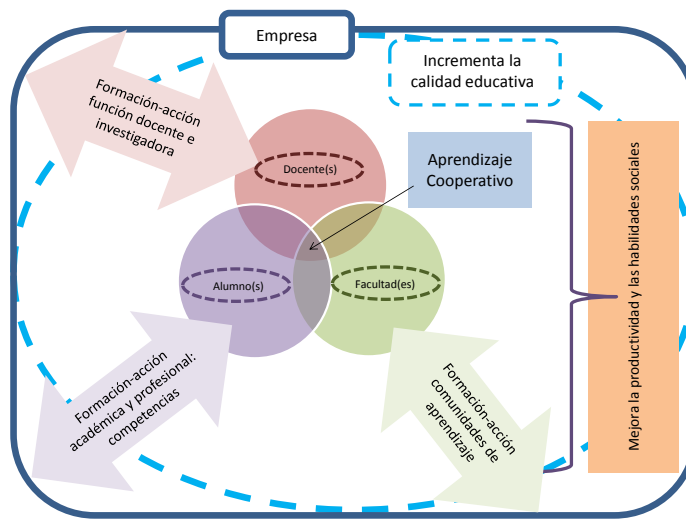
### ***Implicaciones educativas***

Dentro de las implicaciones educativas, se hace necesario puntualizar algunos aspectos relativos a las propias relaciones que se producen al aplicar la metodología de aprendizaje cooperativo en el ámbito universitario.

Al procurar esta metodología la mejora de la productividad y las habilidades sociales para el trabajo y, pretender un contexto favorable al aprendizaje, tanto los docentes entre sí, como los alumnos y toda la facultad interactuarán positivamente en una formación continua que resonará en el clima y gestión educativa en: las comunidades de aprendizaje (fomentadas por la facultad), las competencias académicas y profesionales (requeridas por el alumnado) y la función docente e investigadora (indispensables para el profesorado).

Sin pretender una visión reduccionista de un tema tan complejo como la calidad educativa, la innovación educativa a través de la incursión del aprendizaje cooperativo trascenderá la díada alumno–profesor para la conformación de comunidades de aprendizaje (ilustración 23).

Ilustración 23: Implicaciones educativas del aprendizaje cooperativo



Fuente: elaboración propia

Si se trata del **profesorado**, es indispensable invertir esfuerzos en la formación de los docentes en metodologías activas, una de ellas el aprendizaje cooperativo. Esta formación afectará en dos vías que se entrelazan en una visión de conjunto pero sin perder lo individual.

La primera, el docente fundamentará su hacer en las mejores prácticas para una enseñanza de calidad, lo que implicará conocer las exigencias de una planificación muy precisa y exhaustiva de su clase cooperativa. De esa manera, deberá explicitar en su diseño instructivo los objetivos didácticos que se pretenden alcanzar, definir el tamaño de los grupos de aprendizaje, la asignación de roles y estudiantes a los grupos, la organización física del aula (que se corresponde también con la infraestructura) y la selección de los materiales para garantizar los elementos básicos del aprendizaje cooperativo.

Es el profesor en su función docente quien explica a sus estudiantes cómo alcanzar los objetivos académicos y sociales que se esperan de la actividad cooperativa en donde los criterios de evaluación atienden tanto al contenido a impartir como aspectos vinculados a la interacción social dentro de los grupos. Su rol de “co-acompañante”, de “co-constructor de saberes”, pasa por brindar apoyo en las tareas, responder dudas y preguntas y servir de modelo para el desarrollo de equipos exitosos.

Es el estudiante quien evalúa su propio aprendizaje, son los otros quienes también valoran el procesamiento y funcionamiento del equipo y es el docente quien ha recorrido junto con sus alumnos los pasos necesarios para el logro de la meta, el que evaluará el rendimiento individual y grupal dentro del grupo según el porcentaje asignado en su guía docente.

La segunda, la difusión de la investigación por y para la enseñanza. Es en la investigación en donde la práctica educativa encuentra las mejores evidencias que sustentan un quehacer fundamentado. Es en la sistematización, en la recopilación de otras experiencias, en la incursión de nuevos objetos de estudio en donde el profesor universitario se forma de manera continua para dominar la disciplina que imparte en contenido pero también en estrategias didácticas, metodologías y recursos para un aula cooperativa.

Adoptar la metodología de aprendizaje cooperativo circunscribiéndola solamente a la díada profesor-alumno y alumno-alumno es limitar sus beneficios.

El profesor, convencido de la necesidad de cambio en las estrategias docentes, fortalecerá cooperativamente su función docente e investigadora con otros colegas.

Es en los equipos docentes, donde potenciará la productividad y las habilidades sociales y podrá demandarle a sus alumnos –porque tiene la autoridad moral para hacerlo– capacidad de reflexión y análisis dentro de los grupos de aprendizaje. Pretender exigir a los estudiantes que trabajen cooperativamente cuando son los mismos profesores quienes invierten un gran tiempo de su gestión y aprendizaje profesional de manera individual, es perder una oportunidad valiosísima para seguir co-mejorando la práctica educativa.

Si observamos las implicaciones educativas desde el **alumnado**, se requiere por parte de la universidad –y todavía mucho mas ambicioso del sistema educativo– y del profesorado una sensibilización hacia el trabajo en grupo. ¿Por qué? Dos razones son relevantes en este sentido.

Primero porque la mayoría de los colegios hace a sus alumnos “trabajar en grupo” pero no bajo la modalidad de aprendizaje cooperativo. Y segundo, porque el sistema educativo en el bachillerato español (los dos años previos a los estudios universitarios), alejan aun mas al alumno del trabajo en grupo para centrarse en las pruebas de

selectividad o de acceso al sistema universitario que, por su diseño, obliga a un estudio individualizado.

Resulta llamativo que mientras en los niveles iniciales del sistema educativo, se potencia el trabajo en grupo, a medida que ese mismo estudiante se va acercando al final de su formación escolar, más se ha ido alejando del trabajo grupal, lo que se convierte en una tarea más ardua la inclusión del aprendizaje cooperativo en la universidad.

La implementación de esta metodología en los primeros años de universidad ayuda al alumno a su adaptación porque disminuye las tasas de abandono de las aulas. Y, si se trata de los últimos años, hace que el estudiante comprenda que las formas grupales de trabajo serán un modelo que le acompañará a lo largo de su vida profesional.

La diada alumno–alumno es otra ganancia. Los estudiantes aprenden de otros estudiantes mas o menos expertos. Son los compañeros de clase y, por qué no de otros cursos y disciplinas, una fuente de conocimiento y de intercambio de ideas. Es en la discusión entre pares en donde el conflicto constructivo emerge con mayor fuerza y se potencian los beneficios académicos y sociales como el liderazgo, la comunicación y la toma de decisiones.

Los departamentos, coordinaciones y **facultades** de la universidad entre si y con otras instituciones educativas podrían trabajar en círculos cooperativos concéntricos bidireccionales. Los distintos niveles se nutren de la formación–acción del profesor–aprendiz, del profesor–docente, del profesor–integrante de un equipo docente y, en un nivel macro, de las redes que se establecen entre los equipos docentes–departamento–facultades. Son las comunidades de aprendizaje fundadas en la cooperación las que crean redes de co–conocimiento y de co–investigación que repercutirán a la vez en el ámbito académico y profesional.

Finalmente, la **metodología**, es un calidoscopio de oportunidades según el objetivo planificado y las competencias a desarrollar.

El aula será espacio fundamental para que el aprendizaje cooperativo funcione adecuadamente. Un mobiliario que permita agrupar y reagrupar a todo el salón de clases, la posibilidad de configurar y reconfigurar los espacios para el logro de la interdependencia positiva así como el apoyo de la tecnología no sólo dentro del aula, sino a través del uso de los recursos contemporáneos de las aulas virtuales y sus muchas

posibilidades de comunicación virtual, harán que sea una metodología que pueda volcarse de manera positiva en la universidad.

El aprendizaje cooperativo no se riñe con otras metodologías de enseñanza. Al contrario, la combinación de aprendizajes individuales, cooperativos y competitivos, sin dejar a un lado el uso de la clase magistral, hacen que el proceso de enseñanza sea dinámico, holístico y variopinto, como la vida misma, como son y serán las experiencias académicas y profesionales. Así, la combinación del aprendizaje cooperativo con otras estrategias como el aprendizaje basado en problemas, el estudio de casos y la metodología por proyectos potencian su efectividad.

Sin embargo, su puesta en práctica no está libre de dificultades y pretenderlo sería un absurdo.

Dentro de ese orden de ideas, está el hecho de que muchos estudiantes no son afectos a trabajar con otros compañeros a quienes ellos no elijan. Las consecuencias, bien conocidas por el profesorado, repercuten negativamente en un etcétera como la motivación, el clima del aula, la productividad, la toma de decisiones y en especial, en la evaluación. Aunque “la solución mas sencilla” es dividir al grupo y volver a estructurarlo, puede ser un reto interesante aprender a trabajar bajo esas condiciones... de hecho, así es el ámbito laboral.

Otro problema del que se quejan con regularidad profesores y estudiantes para no trabajar con esta metodología, es que la falta a clase de alguno de los miembros hace que la dinámica grupal se vea afectada y, como es sabido, este tipo de metodología activa se funda en la interdependencia positiva, lo que obliga a los grupos de base cooperativos a convertirse en la plataforma para la tutoría entre iguales.

Son estos grupos de base donde los estudiantes comparten y resuelven dudas y por supuesto, exigen el cumplimiento de las tareas y actividades asignadas a quienes no han asistido. Ante esta eventualidad también compensa la conformación de grupos de tres estudiantes porque en caso de faltar alguno la díada podrá seguir trabajando. Otra posibilidad es la conformación de grupos cooperativos de breve duración que trabajen juntos solo en una sesión (aprendizaje informal) o por periodos breves.

El tamaño de los grupos de clase, es una de las principales razones junto con el tiempo para cubrir todo el temario y la evaluación, motivos para rechazar la implementación de esta metodología en las aulas universitarias.



Recordemos que un buen diseño instructivo guiará la división de la clase en pequeños grupos de trabajo con tareas muy específicas en contenido, dinámica y duración en donde el “aprehender” según los principios del aprendizaje cooperativo favorecerá el control del grupo.

Si se planifican pruebas cortas para ser evaluadas por los propios estudiantes o por los compañeros antes y/o después de la actividad cooperativa, el volumen de corrección para el profesor disminuirá. Algunos autores sugieren asignar un 95% para la evaluación individual y un 5% para el trabajo en grupo y otros, que el porcentaje asignado al trabajo en grupo sea un valor añadido al rendimiento individual. La evaluación debe contemplar la autoevaluación, la co-evaluación y la heteroevaluación.

Los resultados obtenidos en este meta-análisis podrían orientar un cambio en las políticas educativas universitarias para una toma de decisiones fundada en la *Evidencia Basada en la Investigación Educativa*.

## CAPÍTULO VII. LIMITACIONES Y PROSPECTIVA

### *Limitaciones*

Un inconveniente que se ha presentado en la revisión de la literatura ha sido el haber utilizado los **descriptores** *Cooperative Learning* y *Collaborative Learning* porque se han dejado fuera otros asociados a la temática de estudio como *Small Group*, *Learning Together* o nombres específicos de métodos cooperativos como *Peer-Tutoring*, *Computer-Supported Collaborative Learning* (CSCL), *Jigsaw*, *Student Teams-Achievement Divisions*, entre otros.

A esto se suma la dificultad para **diferenciar términos utilizados en ocasiones indistintamente como sinónimos de cooperación o de colaboración** pero que son distintos en su esencia, a saber *group*, *team*, *work groups*, *teamwork*, *group size* porque en estas agrupaciones

students working individually on different sections of a group assignment and then ‘sticking’ them together would not be said to have learned collaboratively. Though group members would still interact, their interactions would not focus on content but on managing the process of determining, allocating and then combining the chunks of work to be completed individually (Summers & Volet, 2010, p. 474).

No siendo una limitación relacionada directamente con el método empleado pero sí vinculada con el proceso de la revisión de la literatura, el **no haber combinado los términos *Cooperative Learning* y *Collaborative Learning* con otros propios de la universidad como *university*, *college*, *undergraduated* y *higher education*** por no aparecer siempre en los títulos de las investigaciones o en las palabras claves – específicamente en la base de datos *Web of Knowledge*–, ha significado por un lado, la ventaja (a) de no reducir dramáticamente el número de estudios registrados y (b) una revisión mas rigurosa para una primera identificación de investigaciones universitarias y por el otro, (c) una tarea muy laboriosa, en donde muy probablemente se hayan excluido estudios debido a la omisión involuntaria.

**La falta de información estadística en los estudios primarios** para realizar los análisis necesarios para el cálculo de la magnitud del efecto ha sido una limitación que atañe directamente a este tipo de síntesis cuantitativa “A common problem in research synthesis is missing information for computing an effect size” (Pigott, 2009, p. 400). No

todos los estudios primarios han incluido las puntuaciones directas necesarias para el cálculo del efecto (tamaño de la muestra, medias y desviaciones estándar...) lo que ha significado la utilización de otros estadísticos por ejemplo *t*-test, el valor *F* del ANOVA o *p*-valores.

Al respecto, Lipsey y Wilson (2001) consideran que las diferentes fórmulas representan grados de aproximación al valor de la magnitud del efecto. De esta forma, las fórmulas mejor valoradas son aquellas que utilizan los datos en bruto (medias y desviaciones) seguidas de fórmulas algebraicamente equivalentes (*t*-test) y en un menor grado, aquellas que refieren al valor exacto de probabilidad para una prueba *t* de Student y las aproximaciones basadas en los datos continuos (coeficiente de correlación). Luego, en un segundo nivel, se encontrarían las estimaciones de las diferencias de media (medias ajustadas, *gain score*...) y estimaciones de la desviación estándar agrupada (ANOVA, ANCOVA...)

El no compartir las puntuaciones directas, es una realidad a la que se ven sometidos quienes llevan a cabo estudios meta-analíticos, lo que revela el no cumplimiento del código ético de organizaciones como la American Psychological Association (2001, 2009) que exige que los datos deben estar a disposición de la comunidad científica e incluso en el caso de efectos insignificantes para permitir así una “síntesis imparcial”. Esta denuncia –que no es reciente– continúa siendo debate en revistas científicas aunque sin grandes cambios (Botella y Ortego, 2010).

A esto se suma el no haber obtenido respuesta por parte de los autores para conseguir estudios no publicados –lo que se llama en inglés *file drawer problem*– o la dificultad propia de encontrar resultados no significativos porque igualmente tienden a no ser difundidos. Estos inconvenientes son propios del método aunque lamentablemente suelen ser muy comunes en la mayoría de los meta-análisis y el presente estudio tampoco escapa a esta dificultad (Wicherts, Borsboom, Kats & Molenaar, 2006).

En consecuencia, existe la posibilidad de que las magnitudes del tamaño del efecto se hayan sobreestimado debido a la falta de esta información aunque también se ha demostrado que no hay evidencia de sesgo de publicación a través del test de Egger y colaboradores (1997) y del número de seguridad de Orwin (1983).

A esta limitación propia del método se añade otra: **la heterogeneidad de los estudios.**

La heterogeneidad en la metodología, diseño, operativización de las variables y conceptualización de los diferentes estudios puede dificultar enormemente su combinación. A este respecto se deberán tomar y clarificar múltiples decisiones que permitan homogeneizar la codificación de las características y resultados de los estudios, transformándolos a una métrica común que los haga directamente comparables (Marín, Sánchez-Meca y López, 2009, p.113).

Y así ha sido. Se ha atendido esta limitación a través de estrategias para la codificación y estimación de los efectos según la naturaleza de las variables de los estudios individuales.

Si bien **la concordancia entre observadores indica un acuerdo aceptable pero no adecuado ni excelente** en la valoración de la calidad metodológica de los estudios (Landis & Koch, 1977), no deja de ser cierto que el nivel de acuerdo alcanzado y los escasos 15 estudios seleccionados al azar que sirvieron para su evaluación tendrían que mejorarse. Sabiendo que el meta-análisis se trata de un método que requiere preferiblemente de un equipo multidisciplinar por su complejidad, en especial para la toma de decisiones en fases muy específicas, el investigador junto con otro observador han trabajado conjuntamente en dos momentos. El primero, en la prueba piloto del manual de codificación. El segundo, en la valoración de la calidad metodológica, ambos con la finalidad de minimizar los sesgos porque

In any meta-analysis must participate an expert on the substantive field of inquire under study. When carrying out a meta-analysis it is necessary to make many critical choices for which the better basis is a deep knowledge of the field from inside. However, it is also advisable to work together with someone with experience in research methodology and statistical analysis (Botella & Gambara, 2006, p. 436).

Estando al tanto de la relación complementaria entre la variabilidad entre grupos ( $Q_b$ ) y la variabilidad intragrupos ( $Q_w$ ), **no se han encontrado variables moderadoras en donde las diferencias entre grupos hayan sido estadísticamente significativas y a la vez, una variabilidad intragrupos no significativa**. Lo esperado habría sido encontrar este supuesto porque entonces se podría afirmar que la variable moderadora por la que se ha agrupado a los estudios, explicaría un alto porcentaje de la heterogeneidad de la magnitud del efecto (Huedo-Medina y Johnson, 2010).

**La falta de información relativa al aprendizaje cooperativo como objeto de estudio** ha sido otra dificultad que ha incidido directamente en el estudio de su propia definición y operacionalización.

Un número importante de documentos no menciona el número de integrantes, la duración del programa, la conformación de los grupos, los métodos cooperativos empleados, los roles asignados y las condiciones que atañen directamente al funcionamiento de los grupos como el sistema de recompensa, estructura de la tarea o la igualdad de oportunidades para la puntuación.

Otros asumen como métodos cooperativos agrupaciones que son en realidad “pseudométodos” (Johnson y Johnson, 1999c) que aunque cumplen con algunos de los principios del aprendizaje cooperativo no siguen con la clasificación y metodología propuesta por los expertos. Esta situación se agrava cuando los métodos empleados utilizan la tecnología como medio para el aprendizaje colaborativo en modalidad presencial o virtual. El crecimiento exponencial propio del área aunado a un tipo de aprendizaje que se caracteriza por ser poco estructurado y en ocasiones hasta informal dificulta aun mas encontrar las similitudes entre los estudios.

Por consiguiente, el número de estudios primarios analizados en estas variables es en ocasiones pequeño (en algunos casos no sobrepasa el valor de 30 estudios) lo que ha generado **una baja potencia estadística** por lo que los resultados obtenidos tienen que ser interpretados con reserva (Hedges & Pigott, 2004). Los estudios pequeños tienden a producir magnitudes del efecto mayores e inestables (Levine, Asada & Carpenter, 2009).

El hecho de que no se especifiquen características tan vitales sigue acotando la comprensión del funcionamiento de los grupos cooperativos, la trascendencia a posibles réplicas de esos estudios primarios y posteriores síntesis cuantitativas.

**Esta dificultad también se ha presentado en la definición y operacionalización del grupo control.** Por ejemplo, algunos documentos primarios asumen de manera indistinta el grupo control como *tradicional learning* y otros como *individual learning*.

En la definición de la enseñanza tradicional, para Lou, Abrami, Spence, Poulsen, Chambers y D'Apollonia (1996) el énfasis recae en un único grupo de estudiantes a quienes un profesor que se encuentra delante de la clase enseña de una manera uniforme a todo el aula. Mientras que para Doymuş (2008b) en el aprendizaje individual, los temas que los estudiantes necesitan aprender requieren de informes y evaluaciones individuales.

La diferencia entre ambas definiciones conceptuales podría ser otra fuente mas de heterogeneidad a explorar en otras síntesis integrativas aunque hasta ahora todas las revisiones cuantitativas consultadas en aprendizaje cooperativo han asumido como “similares” ambas condiciones de la misma forma que se ha asumido en el presente meta-análisis.

*A partir de los hallazgos encontrados ¿Hacia dónde se dirige la investigación?*

Los resultados encontrados son fuente de nuevas interrogantes porque **preguntar una y otra vez permitirá introducir al investigador en un nuevo objeto de estudio por explorar.**

Se recomienda **la ampliación de la muestra** del presente meta-análisis con la incorporación de estudios primarios posteriores al año 2012, la revisión de otras bases de datos, en especial de revistas especializadas en Europa y en asignaturas menos exploradas como Agricultura, Economía, Ingeniería, Física y Enfermería. Son estas dos últimas las que parecieran prometer una eficacia interesante cuando se implementa en las aulas el aprendizaje cooperativo.

También **se sugiere trabajar con el término *Collaborative Learning* en la base de datos *Web of Knowledge*** porque ha sido la única fuente primaria de las consultadas en donde no se ha utilizado este descriptor, debido al extenso volumen de estudios encontrados, no siendo factible su análisis por razones de tiempo que escapan a los límites de este trabajo.

**Además de este descriptor se proponen otros relacionados con métodos cooperativos específicos** como *Scripted Cooperation*, *Jigsaw*, *Reciprocal Peer Tutoring*, *Student Teams–Achievement Divisions*, *Teams–Games–Tournament*, entre otros, los cuales servirán para ampliar la revisión de la literatura y así incrementar el número de posibles estudios primarios.

A este reto, se une el **profundizar en variables moderadoras que afectan el funcionamiento de los grupos cooperativos** que aunque han obtenido un tamaño moderado nada despreciable ninguna sobresale con un efecto grande.

Tal es el caso de los métodos cooperativos que requieren de una especialización de la tarea, en donde no se ofrecen recompensas o gratificaciones al grupo y cuando no todos los miembros del grupo disponen de las mismas oportunidades para contribuir con su desempeño a la puntuación del grupo lo que se traduce en calificaciones diferentes para cada integrante.

De igual forma, otras variables que serían interesantes seguir indagando refieren a las áreas en que se imparte esta metodología de aprendizaje. Es el caso de Ciencias, Ciencias de la Salud y Educación, en programas de baja duración –entre 1 y 4 semanas–, con grupos heterogéneos impares o conformados entre 4 y 6 miembros y con métodos cooperativos que se apoyan en el uso de la tecnología *Computer Support*.

**Las Tecnologías de Información y Comunicación en situaciones cooperativas/colaborativas son escenarios exponenciales de interacción para continuar explorando.** Tan es así que existen revistas especializadas desde el año 1992 como *Computer Supported Cooperative Work: The Journal of Collaborative Computing* y otras mas recientes, creadas en el año 2006 como el caso de *International Journal of Computer-Supported Collaborative Learning* (CSCW) con la finalidad de debatir sobre cuestiones teóricas, prácticas, técnicas y sociales de innovación en y para la colaboración con el apoyo del ordenador.

**Otro valor añadido a incorporar en las futuras síntesis cuantitativas es el curso académico en que se ha trabajado el aprendizaje cooperativo.** Si son estudiantes que inician sus estudios o si esta metodología se ha implementado al final de sus carreras ofrecerá información valiosa en cuanto a su eficacia en el rendimiento. Es sabido que esta metodología disminuye el fracaso escolar y mejora la cohesión grupal, en especial en los primeros cursos. Hasta el momento, solo el meta-análisis de Springer, Stanne y Donovan (1999) lo ha contemplado obteniendo por ejemplo un efecto moderado nada desalentador para el primer curso ( $d=0,52$ ).

**Continuar con el incipiente análisis bibliométrico que se ha comenzado** como disciplina que hoy en día se ha insertado en el ámbito de los *Information Science Studies*. Proseguir con este análisis como objetivo general de una línea de investigación contribuirá a una descripción y evaluación de la actividad científica y sus actores (Jiménez-Contreras, 2000) en cuanto a la producción científica del aprendizaje cooperativo en el ámbito universitario.

Compartiendo el mismo interés y preocupación de Colliver, Feltovich y Verhulst (2003) en la búsqueda de las condiciones mas eficaces para la implementación del aprendizaje cooperativo en el ámbito universitario, se sugiere **contrastar las magnitudes del efecto de los estudios individuales entre las distintas síntesis cuantitativas** y los hallazgos encontrados. Un análisis comparativo de las condiciones sustantivas y metodológicas y de los factores que podrían incidir en el cálculo de esta medida invitarían a discutir el por qué estudios mas actuales han obtenido efectos mas grandes en contraste con estudios menos recientes.

**La cultura de procedencia de las investigaciones y de los estudiantes que participan en ella es otra variable atractiva** que podría examinarse. Los efectos del aprendizaje cooperativo en la productividad no escapan de la cultura en donde se encuentra inmerso este tipo de aprendizaje.

Autores como Thanh, Gillies y Renshaw (2008) luego de la revisión narrativa de estudios solo con estudiantes universitarios asiáticos explican el por qué en esa cultura esta metodología no resulta mas efectiva que el aprendizaje competitivo y el individual. Para estos autores en las culturas gregarias –en inglés *collectivistic cultures*– en donde el profesor es considerado un gurú, el aprendizaje cooperativo no es bien acogido porque aunque la dinámica cultural es colectiva, necesita un aprendizaje estructurado para el control de la incertidumbre.

Otros como Kyndt et al. (2013) llegan a hacer una distinción muy marcada entre la cultura individualista de occidente y la cultura gregaria de oriente. En una clase donde predomina una cultura colectiva, la armonía y el buen clima del aula es muy importante y los conflictos tienen que resolverse. En cambio, en una cultura individualista, lo individual toma el lugar central y las relaciones entre los individuos se pierden.

Al respecto Cox, Lobel y McLeod (1991) hasta llegan a comentar –no dejando de ser necesario matizar esta afirmación para no caer en generalizaciones– que “cross-cultural studies have shown that Northern and Western Europeans and North Americans tend to be individualistic and that Chinese people, other Asians, Latin-Americans, and most East and West Africans tend to be collectivists” (p. 828).

**Continuar mejorando la escala para la valoración de la calidad metodológica de los estudios primarios** es otra tarea pendiente, dado su interés permanente en la comunidad científica por sus implicaciones en la investigación basada en la evidencia



(Altman, Moher & Schulz, 2012; Wicherts, Bakker & Molenaar, 2011) convirtiéndose en el objeto de futuras tesis doctorales.

La mayoría de las síntesis cuantitativas consultadas e inclusive el presente estudio evalúan la efectividad de los grupos cooperativos formales e informales en donde se trabajan los métodos cooperativos estudiados quedando fuera otros tipos de aprendizaje cooperativo como **los grupos cooperativos de base** (Johnson & Johnson, 1999e; Johnson, Johnson & Holubec, 1999). La tarea de incrementar en la universidad el número de estudios centrados en este tipo de aprendizaje cooperativo y su efectividad en el rendimiento y en las habilidades sociales, es un nuevo desafío pero encontrar estudios primarios con este tipo de intervenciones es un reto aun mayor para las revisiones cuantitativas.

La incursión en **síntesis sistemáticas de investigación cualitativa** en aprendizaje cooperativo es una línea futura prometedora.

Sería una aportación partir de la base de datos que se ha creado para la revisión de la presente literatura, y seleccionar solo aquellos estudios cualitativos para un posterior análisis de contenido describiendo la frecuencia en que aparecen categorías relativas a la dinámica de los grupos cooperativos a favor o en contra de la eficacia en el rendimiento académico *–meta–summary–*, o solo elegir los estudios cualitativos con la finalidad de interpretar y refinar significados y conceptos vinculados a la teoría *–meta–synthesis–* o por qué no, un análisis solo de los métodos de investigación empleados por los investigadores en estos estudios ya registrados *–meta–study–* (Saini & Shlonsky, 2012).

Por último, **se aconseja realizar un meta–análisis acumulado** en donde se podrían incorporar nuevos estudios en las variables ya estudiadas con la finalidad de ofrecer una evolución del estimador ponderado según la variable ordinal añadida. Es un trabajo a mediano y largo plazo que permitirá medir en el tiempo la contribución de nuevos estudios al presente meta–análisis.

Igualmente **una meta–regresión es otra propuesta** porque de hacerse este análisis se reconducirán los hallazgos encontrados a un modelo mas complejo en donde las características de los estudios mas relevantes podrían explicar el porcentaje de varianza de los tamaños del efecto. Por ejemplo, las variables moderadoras que formarían parte

del modelo explicativo se reducirían principalmente a las áreas de conocimiento, los métodos cooperativos y los años de publicación.

*Más allá de los resultados encontrados ¿Hacia dónde se dirigen las investigaciones en aprendizaje cooperativo?*

Una mirada general del camino que ha comenzado a recorrer el aprendizaje cooperativo y que seguirá transitando en el interés científico en el corto y mediano plazo es el estudio de esas variables influyentes que esclarecen las condiciones de su eficacia en el rendimiento académico de estudiantes universitarios.

Autores como Coll en 1991 (citado por Gavilán y Alario, 2010) ya afirmaba que solo el 20% de la varianza total es explicada a través de algunas esas variables como son la tarea, la edad, la comunicación, la repetición de lo aprendido, entre otras, en donde las investigaciones futuras tendrán que continuar profundizando en esas condiciones porque el 80% restante queda aun por indagar. Esto significa que muchas otras son las variables que explican el por qué de la superioridad de la cooperación como situación de aprendizaje cuando se compara con el aprendizaje competitivo e individual.

Por tanto, los factores que justifican una mayor investigación en los próximos años son aquellos asociados con los **mecanismos mediadores implicados en el proceso interactivo** de los grupos como la comunicación y su variabilidad en función del contenido y el método cooperativo y, los relacionados con los procesos cognitivos y de regulación del comportamiento (León del Barco, Felipe, Iglesias y Marugán, 2014).

Estos procesos orientados a la acción, implican el estudio riguroso de la **interdependencia positiva y la responsabilidad individual** (Johnson & Johnson, 2014b) pero también de la coordinación de esas actuaciones emergentes.

Al respecto, hay autores que proponen no el estudio de la coordinación explícita, es decir, de las estrategias y patrones de conducta que se emplean cuando los estudiantes planifican intencionalmente sus acciones, sino en **examinar la coordinación implícita en donde un individuo del grupo se anticipa a la necesidad de los otros sin que esto haya requerido de una comunicación y planificación manifiesta** (Gil, Rico y Sánchez-Manzanares, 2008; Rico, Sánchez-Manzanares, Gil & Gibson, 2008).

No es de extrañar que Serrano y Pons (2014) coincidan en señalar que son los mecanismos que surgen en la interacción intragrupal e intergrupala los que apuntan más alto en la eficacia de los grupos. La no atención y matización de estas interacciones podría ser uno de los motivos de los resultados negativos y contradictorios en la implementación de esta metodología en las aulas universitarias.

Otros factores que también han sido estudiados en menor escala y que, en consecuencia comienzan a interesar refieren a las **condiciones previas que subordinan la eficacia del aprendizaje cooperativo** como son las características personales y cognitivas. Son un ejemplo, el conocimiento previo (Van Blankenstein, Dolmas, Van der Vleuten & Schmidt, 2013), los estilos de personalidad, las habilidades cognitivas de inducción y extraversión y la orientación social, las cuales se asocian positivamente con un aprendizaje cooperativo efectivo (León del Barco, Felipe, Iglesias y Marugán, 2014).

A este listado se suma la tendencia en el estudio de los **procesos motivacionales y afectivos** como la cohesión grupal y los estados de ánimo, emociones y sentimientos compartidos. Se está comenzado a hilar fino en el estudio de variables que tienen que ver con la “eficacia colectiva –eficacia percibida para realizar exitosamente una tarea– y el potencial grupal –eficacia percibida para afrontar con éxito cualquier tarea–” (Gil, Rico y Sánchez–Manzanares, 2008, p. 29) pero en especial, en el conflicto que emerge en la interacción (Johnson & Johnson, 2014b).

Investigaciones centradas en el análisis del discurso de las conversaciones que afloran entre los miembros del grupo, es una línea de investigación comenzada hace años por Neil Mercer (1997, 2010) y que encuentra cada vez más un mayor número de discípulos. **Es el conflicto** que emerge de la tarea y no el que surge en las relaciones entre los miembros del grupo un factor que en la mayoría de los casos favorece el aprendizaje y que continúa siendo estudiado. Es la **controversia constructiva** en donde esas discrepancias intentan ser superadas a través de acuerdos cuando existen diferencias de información, ideas, puntos de vista y opiniones (Johnson, Johnson & Tjosvold, 2008).

El **trabajo cooperativo entre profesores y entre profesores y facultad** son líneas de actuación hoy en día muy prometedoras. Pensar en equipos docentes que atienden su función docente–investigadora y que transforman el lugar de trabajo para y con los

otros, en una comunidad de aprendizaje, es una fuente de conocimiento inagotable asociada a la productividad.

Los marcos teóricos, metodológicos y de análisis de grupos de “aprendizaje no formalizado” entre profesores están en plena construcción en España (Fernández, Guisasola, Garmendia, Alkorta y Madinabeitia, 2013; Mauri, Clarà, Ginesta y Colomina, 2014) y fuera de España (Bakkenes, Vermunt & Wubbels, 2010; Boud & Middleton, 2003; Tynjälää, 2008) pero también, se están levantando los cimientos en el estudio del “aprendizaje formal” de profesores universitarios españoles, quienes por ejemplo, han recibido un programa en aprendizaje cooperativo para la organización de aulas cooperativas según el enfoque de competencias (Serrano y Pons, 2014).

Actualmente, existen revisiones narrativas y meta-análisis que se nutren de los principios básicos de la cooperación pero que se centran, no en el aprendizaje cooperativo como metodología de enseñanza-aprendizaje con sus métodos y variantes, sino en la eficacia de los grupos, llamada en inglés *team effectiveness*.

Son Kozlowski e Ilgen (2006) quienes aseguran que esas revisiones están dejando fuera dos tendencias emergentes que se están convirtiendo en un tema a profundizar “these two trends are the increasing proliferation of virtual teams and the growing use of teams composed of diverse, multiethnic, and multicultural members” (p.114). Es la confluencia entre **lo virtual y los equipos multiculturales** una fuente inexcusable de investigación. A lo que agregan

And, while the research need is often framed as how to mitigate the negative effects of virtuality and multicultural diversity on team processes, we think the greater theory and research challenge is how to harness the emerging technological capabilities to enhance and evolve team processes in virtual environments that cut across different cultures (p.114).

Estos estudios ya están **trascendiendo el contexto educativo para incursionar en organizaciones y empresas** (Gupta, Mattarelli, Seshasai & Broschak, 2009; Reimer, Reimer & Czienskowski, 2010; Richter, Dawson & West, 2011) con respecto al liderazgo (Wang, Waldman & Zhang, 2014) en donde se evalúan intervenciones para la mejora de la productividad (Salas, Nichols & Driskell, 2007) y la efectividad en equipos virtuales (De Guinea, Webster & Staples, 2012; Mesmer-Magnus, DeChurch, Jimenez-Rodriguez, Wildman & Shuffler, 2011).

**La combinación de técnicas cuantitativas y cualitativas que permitan evaluar la eficacia de los grupos cooperativos** es una apuesta metodológica que va ganando terreno en el campo de la medición (MacPherson, Jones, Whitehouse & O'Neill, 2001; Willis et al. 2002).

El enfoque mixto o los métodos mixtos de investigación parten de la premisa central de que el uso de la combinación de enfoques ofrece una mejor comprensión de la situación de estudio cuando se compara con la aplicación de métodos individuales (Creswell & Plano, 2007; Tashakkori & Teddlie, 2011).

Son una muestra de ello, ejemplos recientes de estudios primarios para comprender y profundizar en variables moderadoras de la cooperación, el empleo de técnicas y métodos cualitativos como la observación, la entrevista a profundidad, el auto-reporte (Summers & Volet, 2010), los grupos focales (Skala, Slater & Adams, 2000), la etnografía virtual (García y Suárez, 2011) y la teoría fundamentada (Singaram, Van der Vleuten, Stevens & Dolmans, 2011; Sorensen & Bernard, 2012).

También es cierto, que **las recomendaciones metodológicas siguen siendo una línea de alta producción investigadora** que trascienden el objeto de estudio de esta tesis. Las críticas constructivas atañen a los métodos y análisis estadísticos empleados por los investigadores con la finalidad de afinar aun mas los resultados encontrados, lo que sin duda está repercutiendo y seguirá repercutiendo en los futuros estudios meta-analíticos.

Es precisamente, el **cálculo de la magnitud del efecto** una medida que comienza a exigirse desde hace un tiempo pero cada vez con mayor fuerza con el propósito de no sustituir sino de complementar el  $p$ -valor y el contraste de hipótesis porque de esta forma permite “to get better statistics and results for a long time” (American Educational Research Association, 2006; American Psychological Association, 2001; Anderson, Burnham, & Thompson, 2000; Kirk, 1996; Plucker, 1997; Robinson & Levin 1997; Thompson & Snyder, 1997, todos estos autores y organizaciones citados en Sun, 2008, p. 1).

Los beneficios de este estadístico se resumen en implicaciones para la toma de decisiones porque ofrecerá a otros investigadores comprender la magnitud actual del efecto de la intervención de interés. Al tratarse de una “metric-free measures of the size of mean differences or the strength of relations, they may be used to compare the results of different studies to one another” (Sun, 2008, p.5) permitirá que “before a study is

carried out, estimates of anticipated effect sizes can be used to project the sample size that would be adequate for detecting statistically significant results” (Sun, 2008, p. 5).

Son estas breves páginas una pequeña muestra de que el estudio del grupo como fenómeno social sigue avanzando porque “There is a solid foundation for concluding that there is an emerging science of team effectiveness” (Kozlowski & Ilgen, 2006, p.77).

*De modo que mis respuestas a las preguntas “¿Cómo lo sabe? ¿Cuál es la fuente o la base de su afirmación? ¿Qué observaciones lo han conducido a ella? Sería: “Yo no lo sé; mi afirmación era meramente una presunción. No importa la fuente, o las fuentes, de donde pueda haber surgido. Hay muchas fuentes posibles y yo quizás no conozca ni la mitad de ellas; en todo caso, los orígenes y las genealogías son poco atinentes al problema de la verdad. Pero si usted está interesado en el problema que yo trato de resolver mediante mi afirmación tentativa, puede usted ayudarme criticándola lo más severamente posible que pueda; y puede refutar mi afirmación, lo ayudaré gustosamente, en todo lo que de mí dependa, a refutarla*  
*Karl Popper, 1972, traducción de Néstor Miguez 1991, p. 51*



## Referencias

Los estudios que han sido seleccionados para el meta-análisis se han indicado con un asterisco

- \*Abdul, S. (2009). Cooperative learning versus the lecture method of instruction in an introductory statistics course. *Jurnal Sains dan Matematik*, 1(1), 59–71.
- Abrami, P., Lou, Y., Chambers, B., Poulsen, C. & Spence, J. (2000). Why should we group students within-class for learning? *Educational Research and Evaluation*, 6(2), 158–179.
- \*AbuSeileek, A. (2007). Cooperative vs. individual learning of oral skills in a CALL environment. *Computer Assisted Language Learning*, 20(5), 493–514.
- Acarturk, C., Cuijpers, P., Van Straten, A. & De Graaf, R. (2009). Psychological treatment of social anxiety disorder: a meta-analysis. *Psychological Medicine*, 39, 241–254.
- ACCENTURE (2007). *Las competencias profesionales en los titulados. Contraste y diálogo Universidad-Empresa*. Recuperado de <http://www.ugr.es/~filosofia/recursos/espacio-europeo/2007-Universia-competencias-profesionales-en-titulados.pdf>
- Adell, J. (Junio, 2011). *Perspectivas teóricas en torno al aprendizaje colaborativo mediado por TIC*. I Congreso de Aprendizaje Colaborativo y TICs (CIMAC). Universidad de Salamanca.
- Agarwal, R. & Nagar, N. (2011). *Cooperative learning*. Delhi: Kalpaz publications.
- Akdemir, E. & Arslan, A. (2012). From past to present: trend analysis of cooperative learning studies. *Procedia-Social and Behavioral Sciences*, 55, 212–217.
- Akhtar, K., Perveen, Q., Kiran, S., Rashid, M. & Khatoon, A. (2012). A Study of Student's Attitudes towards Cooperative Learning. *International Journal of Humanities and Social Science*, 2(11), 141–147.
- Alderson, P. & Green, S. (2009). *Cochrane Collaboration open learning material for reviewers*. Version 1.1 Recuperado de <http://www.cochrane-net.org/openlearning/PDF/Openlearning-full.pdf>
- Alorda, B., Suenaga, K. & Pons, P. (2011). Design and evaluation of a microprocessor course combining three cooperative methods: SLDA, PjBL and CnBL. *Computers and Education*, 57, 1876–1884.



- Altman, D., Moher, D. & Schulz, K. (2012). Improving the reporting of randomised trials: the CONSORT Statement and beyond. *Statistics in Medicine*, 31(25), 2985–2997.
- Álvarez, F. y García, J. (2012). El aprendizaje cooperativo en el Máster de Secundaria para enseñar y aprender a ser tutor. *Revista Educativa Hekademos*, 12, 83-93.
- American Psychological Association (2001). Ethical standards for the reporting and publishing of scientific information. In *Publication Manual of the American Psychological Association* (5ª Ed). Washington, DC: Author.
- American Psychological Association (2009). *Manual of the American Psychological Association* (6ª Ed). México: Manual Moderno.
- Andreu, L., Sanz, M. y Serrat, E. (2009). Una propuesta de renovación metodológica en el marco del Espacio Europeo de Enseñanza Superior: los pequeños grupos de investigación cooperativos. *REIFOP*, 12(3), 111–126.
- Apodaca, P. (2009). Estudio y trabajo en grupo. En M. De Miguel Díaz (Coord.). *Metodologías de enseñanza y aprendizaje para el desarrollo de competencias. orientaciones para el profesorado universitario ante el Espacio Europeo de Educación Superior* (pp.169–190). Madrid: Alianza editorial.
- Archer-Kath, J., Johnson, D. & Johnson, R. (1994). Individual versus group feedback in cooperative groups. *Journal of Social Psychology*, 134(5), 681–694.
- Arco-Tirado, J., Fernández-Martín, F. & Fernández-Balboa, J. (2011). The impact of a peer-tutoring program in quality standards in higher education. *High Education*, 62, 773–788.
- Arnaiz, P. (1989). *El aprendizaje en grupo en el aula*. Barcelona: Graó.
- Aronson, E., & Patnoe, S. (2011). *Cooperation in the classroom: The jigsaw method* (3ª Ed.). London: Pinter & Martin, Ltd.
- Aydin, F. (2013). Opinions and Self-evaluation of the Students of Geography Department about the Applications of Cooperative Learning in Regional Geography Course. *Educational Sciences: Theory & Practice*, 13(4), 2413–2418.
- \*Baghcheghi, N., Reza, H. & Rezaei, K. (2011). A comparison of the cooperative learning and traditional learning methods in theory classes on nursing students' communication skill with patients at clinical settings. *Nurse Education Today*, 1–6.

- \*Bahar-Özvaris, S., Çuhadaroglu, F., Turan, S. & Peters, A. (2006). Cooperative learning: a new application of problem-based learning in mental health training. *Medical Teacher*, 28(6), 553–557.
- \*Baker, L. (1995). *The Effect of Cooperative Study Groups on Achievement of College-level Computer Science Programming Students* (Tesis doctoral). The University of Texas at Austin, USA.
- Bakkenes, I. Vermunt, J. & Wubbels, T. (2010). Teacher learning in the context of educational innovation. Learning activities and learning outcomes of experienced teachers. *Learning and Instruction*, 20, 533–548.
- Ballantine, J. & McCourt, P. (2009). Accounting Undergraduates' Perceptions of Cooperative Learning as a Model for Enhancing their Interpersonal and Communication Skills to Interface Successfully with Professional Accountancy Education and Training. *Accounting Education: an International Journal*, 18(4–5), 387–402.
- Bandura, A. (1989). Social cognitive theory. *Annals of Child Development*, 6, 1–60.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp.71–81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998).
- Recuperado de [http://www.stanford.edu/~kcarmel/CC\\_BehavChange\\_Course/readings/Bandura\\_Selfefficacy\\_1994.htm](http://www.stanford.edu/~kcarmel/CC_BehavChange_Course/readings/Bandura_Selfefficacy_1994.htm)
- Barkley, E., Cross, P. y Howell, C. (2007). *Técnicas de aprendizaje colaborativo*. Madrid: Ediciones Morata.
- \*Basili, P. (1988). Conceptual change strategies within cooperative groups of community college chemistry students: an experiment (Tesis doctoral). University of Maryland, USA.
- \*Bayraktar, G. (2011). The effect of cooperative learning on students' approach to general gymnastics course and academic achievements. *Educational Research and Reviews*, 6(1), 62–71.
- Becerra, C., Gras-Martí, A., Hernández, C., Montoya, J., Osorio, L. y Sancho, T. (2012). Renovación de la Enseñanza Universitaria Basada en Evidencias (REUBE). Una metodología de acción flexible. *Perfiles Educativos*, XXXIV(135), 62–77.

- \*Beck, L. & Chizhik, A. (Marzo, 2008). *An experimental study of cooperative learning in CS1*. Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education, Portland, Oregon, USA.
- Beck, L. & Chizhik, A. (2013). Cooperative Learning Instructional Methods for CS1: Design, Implementation, and Evaluation. *ACM Transactions on Computing Education*, 13(3), 10–1 10–21.
- Benito, A., Bonson, M. e Icarán, E. (2005). Nuevas clases para la docencia universitaria en el Espacio de Educación Superior. En A. Benito y A. Cruz (Eds.) *Metodologías activas* (pp.21–64). Madrid: Narcea S.A.
- \*Bilgin, I. (2009). The effects of guided inquiry instruction incorporating a cooperative learning approach on university students' achievement of acid and bases concepts and attitude toward guided inquiry instruction. *Scientific Research and Essay*, 4(10), 1038–1046.
- Bloom, B. (1976). *Human Characteristics and School Learning*. New York: McGraw-Hill.
- Bloom, D. (2009). Collaborative test taking. Benefits for learning and retention. *College Teaching*, 57(4), 216–220.
- \*Boling, N. & Robinson, D. (1999). Individual study, interactive multimedia, or cooperative learning: which activity best supplements lecture-based distance education? *Journal of Educational Psychology*, 91(1), 169–174.
- Bolívar, A. y Caballero, K. (2008). Cómo hacer visible la excelencia en la enseñanza universitaria. *Revista iberoamericana de Educación*, 46(8), 1–10.
- \*Bölükbaş, F., Keskin, F. & Polat, M. (2011). The effectiveness of cooperative learning on the reading comprehension skills in Turkish as a foreign language. *The Turkish Online Journal of Educational Technology*, 10(4), 330–335.
- Borenstein, M., Hedges, L. & Rothstein, H. (2007). Introduction to meta-analysis. Recuperado de <http://www.meta-analysis.com/downloads/Meta%20Analysis%20Fixed%20vs%20Random%20effects.pdf>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons.
- Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1, 97–111.

- \*Borresen, C. (1990). Success in introductory statistics with small groups. *College Teaching*, 38(1), 26–28.
- Boronat, J. y Ruiz, E. (2011). Alcance de la investigación en el aula en la formación inicial del profesorado. *Tendencias Pedagógicas*, 18, 133–151.
- Botella, J. y Gambara, H. (2002). *Qué es el meta-análisis*. Madrid: Biblioteca nueva.
- Botella, J. & Gambara, H. (2006). Doing and reporting a meta-analysis. *International Journal of Clinical and Health Psychology*, 6(2), 425–440.
- Botella, J. y Ortego, C. (2010). Compartir datos: hacia una investigación más sostenible. *Psicothema*, 22(2), 263–269.
- Boud, D. & Middleton, H. (2003). Learning from others at work: communities of practice and informal learning. *Journal of Workplace Learning*, 15(5), 194–202.
- Bowen, C. (2000). A quantitative literature review of cooperative learning effects on high school and college chemistry achievement. *Journal of Chemical Education*, 77(1), 116–119.
- Bowman, N. (2012). Effect Sizes and Statistical Methods for Meta-Analysis in Higher Education. *Research in Higher Education*, 53, 375–382.
- Boza, A. y Toscano, M. (2012). Motivos, actitudes y estrategias de aprendizaje motivado en alumnos universitarios. *Profesorado. Revista de Curriculum y Formación del Profesorado*, 16(1), 125–142.
- Brewer, S. & Klein, J. (2006). Type of positive interdependence and affiliation motive in an asynchronous, collaborative learning environment. *Educational Technology, Research and Development*, 54, 331–354.
- Bronfenbrenner, U. (1974). The origins of alienation. *Scientific American*, 237(2), 53–61.
- Bruffee, K. (1984). Collaborative learning and the “conversation of man”. *College English*, 46(7), 635–652.
- Bruffee, K. (1995). Sharing Our Toys: Cooperative Learning versus Collaborative Learning. *Change*, 27(1), 12–18.
- Bruffee, K. (1999). *Cooperative Learning. Higher Education, Interdependence, And The Authority Of Knowledge* (2ª Ed.). Baltimore/Londres: The Johns Hopkins University Press.
- Bruner, J. (1998). *Acción, pensamiento y lenguaje*. Madrid: Alianza.
- Bryant, B. (1978). Cooperative Goal Structure and Collaborative Learning. *Teaching of Psychology*, 5(4), 182–185.

- Buchs, C., Pulfrey, C., Gabarrot, F. & Butera, F. (2010). Competitive conflict regulation and informational dependence in peer learning. *European Journal of Social Psychology*, 40, 418–435.
- Bullock, R & Svyantek, D. (1985). Analyzing meta-analysis: potential problems, an unsuccessful replication, and evaluation criteria. *Journal of Applied Psychology*, 70(1), 108–155.
- Bulut, S. (2009). A cross-cultural study on the usage of cooperative learning techniques in graduate level education in five different countries. *Revista Latinoamericana de Psicología*, 42(1), 111–115.
- Callado, F. y Utrero, N. (2012). Características del alumno y aprendizaje cooperativo: aplicación a la enseñanza de la Economía de la Empresa. *@tic Revista D’Innovació Educativa*, 20–25.
- Cambra, C. y Laborda, C. (1998). Función del maestro en la aplicación del aprendizaje cooperativo en un aula de integración con alumnos sordos. *Educar*, 22, 255–299.
- Campbell, D. & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Wadsworth: CENGAGE Learning.
- Campisi, J. & Finn, K. (2011). Does Active Learning Improve Students’ Knowledge of and Attitudes Toward Research Methods? *Journal of College Science Teaching*, 40(4), 38–45.
- Carrier, C. & Sales, G. (1987). Pair versus individual work on the acquisition concepts in a computer-based instructional lesson. *Journal of Computer-Based Instruction* 14(1), 11–17.
- \*Carroll, E. (1989). *A comparative study of achievement in college-level written business communication using lecture and cooperative learning teaching methods* (Tesis doctoral). Oklahoma State University, USA.
- Carugati, F. y Mugny, G. (1988). La teoría del conflicto sociocognitivo. En G. Mugny y J. Pérez (Eds.) *Psicología social del desarrollo cognitivo* (pp.79-188). Barcelona: Anthropos.
- Cassany, D. (2004). Aprendizaje cooperativo para ELE. En *Actas del programa de formación para profesorado de español como lengua extranjera 2003–2004* (pp.11-30). Múnich: Instituto Cervantes de Múnich.
- Catalá, F. y Tobias, A. (2013, en prensa). Metaanálisis de ensayos clínicos aleatorizados, heterogeneidad e intervalos de predicción. *MEDCLI*. Recuperado de <http://dx.doi.org/10.1016/j.medcli.2013.06.013>

- Çelik, S., Aytin, K. & Bayman, E. (2013). Implementing cooperative learning in the language classroom: opinions of Turkish teachers of English. *Procedia–Social and Behavioral Sciences*, 70, 1852–1859.
- Centro Cochrane Iberoamericano (2012). *Manual Cochrane de Revisiones Sistemáticas de Intervenciones, versión 5.1.0* Barcelona. Recuperado de <http://www.cochrane.es/?q=es/node/269>
- Chacón, S., Sanduvete, S. y Alarcón, D. (Septiembre, 2005). *Validez de contenido de una escala para valorar la calidad de los estudios primarios en meta-análisis*. IX Congreso de Metodología de las Ciencias Sociales y de la Salud, Granada, España.
- Chacón, S., Sanduvete, S., Sánchez, M. & Sánchez–Meca, J. (Mayo, 2007). *Enhancing utility of systematic reviews from methodological quality of primary studies. A case study in training programs*. The seventh annual international Campbell Collaboration Colloquium, London.
- Chambers, B., Slavin, R. E., Madden, N. A., Abrami, P., Logan, M. K., & Gifford, R. (2011). Small–group, computer–assisted tutoring to improve reading outcomes for struggling first and second graders. *The Elementary School Journal*, 111(4), 625–640.
- \*Cheng, K. (2009). The comparative effect on business creativity when web based collaborative learning vs. traditional lecturing instruction. *Research in Higher Education Journal*, 2, 1–15.
- \*Cheng, Y. & Ku, H. (2009). An investigation of the effects of reciprocal peer tutoring. *Computers in Human Behavior*, 25, 40–49.
- Cheung, A. & Slavin, R. (2012). How Features of Educational Technology Applications Affect Student Reading Outcomes: A Meta-analysis. *Educational Research Review*, 7(3).
- Cheung, A. & Slavin R. (2013). Effects of Educational Technology Applications on Reading Outcomes for Struggling Readers: A Best-evidence Synthesis. *Reading Research Quarterly*, 48(3), 277-299.
- Chipchase, L., Dalton, M., Williams, M. & Scutter, S. (2004). Is education immune from evidence-based scrutiny? *Australian Journal of Physiotherapy*, 50, 133-135.

- Clarke, A., Trigss, V. & Nielsen, W. (2014). Cooperating Teacher Participation in Teacher Education: A Review of the Literature. *Review of Educational Research*, 84(2), 163–202.
- Clarke, J. (1999). Pieces of the Puzzle: The *Jigsaw* Method. En S. Sharan (Ed.) *Handbook of cooperative learning methods* (pp.34–50). USA: Praeger.
- Clegg, S. (2005). Evidence-based practice in educational research: a critical realist critique of systematic review. *British Journal of Sociology of Education*, 26(3), 415–428.
- \*Clifford, J. (1981). Composing in stages: the effects of a collaborative pedagogy. *Research in the Teaching of English*, 15(1), 37–53.
- Coalition for Evidence-Based Policy (2010). *Checklist for reviewing a randomized controlled trial of a social program or project, to assess whether it produced valid evidence*. Recuperado de <http://coalition4evidence.org/>
- Coe, R. y Merino, C. (2003). Magnitud del efecto: una guía para investigadores y usuarios. *Revista de Psicología de la PUCP*, XXI(1), 146–177.
- Cohen, E. (1994a). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64, 1–35.
- Cohen, E. (1994b). *Designing group work. Strategies for the heterogeneous classroom*. New York: Teacher's College.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2<sup>a</sup> Ed.), Hillsdale, N.J.: Erlbaum. Recuperado de [http://www.lrdc.pitt.edu/schneider/P2465/Readings/Cohen,%201988%20\(Statistical%20Power,%20273-406\).pdf](http://www.lrdc.pitt.edu/schneider/P2465/Readings/Cohen,%201988%20(Statistical%20Power,%20273-406).pdf)
- Cohen, M. (2010). Cooperative Learning in Educational Psychology Modeling Success for Future Teachers. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.69–90). Virginia: Stylus – The National Teaching and Learning Forum.
- Cohen, P., Kulik, J. & Kulik, C. (1982). Educational Outcomes of Tutoring: A Meta-analysis of Findings. *American Educational Research Journal*, 19, 237–248.
- Coll, C. (1984). Estructura grupal, interacción entre alumnos y aprendizaje escolar. *Infancia y Aprendizaje*, 27–28, 119–138.

- Coll, C., Rochera, M. & Gispert, I. (2014). ing online collaborative learning in small groups: Teacher feedback on learning content, academic task and social participation. *Computers & Education*, 75, 53–64.
- Colomina, R. y Onrubia, J. (1990). Interacción educativa y aprendizaje escolar. En C. Coll, A. Marchesi y J. Palacios (Comp.) *Desarrollo psicológico y educación. Psicología de la educación. Vol. II* (pp.415-436). Madrid: Alianza Editorial.
- Colliver, J., Feltovich, P. & Verhulst, S. (2003). Small group learning in medical education: a second look at the Springer, Stanne and Donovan meta-analysis. *Teaching and Learning in Medicine*, 15(1), 2–5.
- Colorado Department of Education (2011). *Fast facts: Evidence-Based Practice*. Recuperado de [http://www.cde.state.co.us/cdesped/download/pdf/FF-EBP\\_MH\\_ADHD.pdf](http://www.cde.state.co.us/cdesped/download/pdf/FF-EBP_MH_ADHD.pdf)
- CONSORT (2010). *Cheklst of information to include when reporting a randomised trial*. Recuperado de <http://www.consort-statement.org/consort-statement/>
- Cook, T. & Campbell, D. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings*. Chicago, Illinois: Rand McNally.
- Cook, T., Campbell, D. & Peracchio, L. (1990). Quasi Experimentation. En Dunnette JD, Hough LM (eds). *Handbook of industrial and organizational psychology* (pp. 491– 573). Palo Alto, CA: Consulting Psychologists Press, Inc. Recuperado de [http://www.owl.net.rice.edu/~antonvillado/courses/09a\\_psyc630001/Cook,%20Campbell%20&%20Peracchio%20\(1990\)%20HIOP.pdf](http://www.owl.net.rice.edu/~antonvillado/courses/09a_psyc630001/Cook,%20Campbell%20&%20Peracchio%20(1990)%20HIOP.pdf)
- Cook, T.D., Cooper, H., Cordray, D.S., Hartmann, H., Light, R.J., Louis, T.A. & Mosteller, F. (1992). *Meta-Analysis for Explanation*. New York: Russell Sage Foundation.
- Cooper, H. (2003). Editorial. *Psychological bulletin* 129, 3–9.
- Cooper, H. & Hedges, L.V. (eds) (1994). *The Handbook of Research Synthesis* (2ª Ed.). New York: Russell Sage Foundation.
- Cooper, H. & Hedges, L. (2009). Potentials and limitations. En H. Cooper, L. Hedges & J. Valentine (Editors). *The handbook of research synthesis and meta-analysis* (2ª Ed.) (pp.562–571). New York: Russell Sage Foundation.
- Cooper, J. (2009). What is cooperative learning. En J. Cooper, P. Robinson & D. Ball (2003). *Small Group Instruction in Higher Education. Lessons from the Past, Visions of the Future*. (2ª Ed.) (pp.3–5). Sitllwater: New Forums.



- Cordero, G. y Serrano, E. (2010). Encuesta retos de la evaluación de los programas de formación de profesores: el caso de un programa de métodos de aprendizaje cooperativo. *Revista Iberoamericana de Evaluación Educativa*, 3(1e), 192–201.
- Corominas, J. y Pascual, J. (2007). *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.
- Cottell, P. (2010). Cooperative Learning in Accounting. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.11–34). Virginia: Stylus – The National Teaching and Learning Forum.
- \*Courtney, D., Courtney, M. & Nicholson, C. (1992). The effect of cooperative learning as an instructional practice at the college level. Paper presented at the annual meeting of the Mid–South Educational Research Association, Knoxville, USA.
- Cox, T., Lobel, S. & McLeod, P. (1991). Effects of ethnic groups’ cultural differences on cooperative and competitive behavior on a group task. *Academy of Management Journal*, 34, 827–847.
- Creswell, J. & Plano, V. (2007). *Designing and conducting mixed methods research* (2<sup>a</sup> Ed.). California: SAGE Publications.
- Crockrell, K. Caplow, J. & Donaldson, J. (2000). A context for learning: collaborative groups in the problem–based learning environment. *Review of Higher Education* 23, 347–364.
- \*Crooks, S., Klein, J., Jones, E. & Dwyer, H. (1996). Effects of Cooperative Learning and Learner–Control Modes in Computer–Based Instruction. *Journal of Research on Computing in Education*, 29(2), 109–123.
- \*Crooks, S., Klein, J., Savenye, W. & Leader, L. (1998). Effects of Cooperative and Individual Learning During Learner–Controlled Computer–Based Instruction. *Journal of Experimental Education*, 66(3), 223–244.
- Cruce, T., Wolniak, G., Seifert, T. & Pascarella, E. (2006). Impacts of good practices on cognitive development, learning orientations, and graduate degree plans during the first year of college. *Journal of College Student Development*, 47, 365–383.
- Cuijpers, P., Van Straten, A., Smits, N. & Smit, F. (2006). Screening and early psychological intervention for depression in schools Systematic review and meta–analysis. *European Child & Adolescent Psychiatry*, 15(5), 300–307.
- Cuseo, J. (1992). Collaborative and cooperative learning in higher education: A proposed taxonomy. *Cooperative Learning and College Teaching*, 2(2), 2–4.

- Cvetkovic, D. (2013). Evaluation of FCS Self and Peer–Assessment Approach Based on Cooperative and Engineering Design Learning. *35<sup>th</sup> Annual International Conference of the IEEE EMBS* (pp.2519-2522). Osaka – Japan. doi: 10.1109/EMBC.2013.6610052
- Damon, W. & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, 13(1), 9–19.
- Davies, P. (2000). The Relevance of Systematic Reviews to Educational Policy and Practice. *Oxford Review of Education*, 26(3&4), 365–378.
- De Arriba, R. y García, M. (2013). Aprendizaje colaborativo en grupos masificados: el uso de la Wiki en el Sistema Económico Mundial. En D. Díaz–Fuentes (Coord.), *Taller de innovación docente. Sociedad de Economía Mundial* (pp.53-61). Santander: Universidad de Cantabria.
- De Backer, L., Van Keer, H. & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students’ metacognitive knowledge and regulation. *Instructional Science*, 40(3), 559–588.
- De Guinea, A., Webster, J. & Staples, S. (2012). A meta–analysis of the consequences of virtualness on team functioning. *Information & Management*, 4, 301–308.
- \*Dees, R. L. (1991). The role of cooperative learning in increasing problem–solving ability in a college remedial course. *Journal for Research in Mathematics Education*, 22, 409–21.
- Dennen, V. & Hoadley, C. (2013). Designing Collaborative Learning through CSCL. En: C. Hmelo–Silver, C. Chinn, C. Chan & O’Donnell, A. *The International Handbook of Collaborative Learning* (pp.389-402). New York: Routledge.
- DerSimonian, R. & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* 7(3), 177-88.
- Des Jarlais, D.C., Lyles, C., Crepaz, N. & the Trend Group (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*, 94, 361–366.
- Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations*, 2, 129–152.
- DeVries, D. (1976). Teams–Games–Tournament: A Gaming Technique That Fosters Learning. *Simulation & Games*, 7(1), 21–33. Recuperado de <http://sag.sagepub.com/content/7/1/21>.

- DeVries, R. (1997). Piaget's social theory. *Educational research*, 26(2), 4–17.
- DeVries, D., Edwards, K. & Slavin, R. (1978). Biracial Learning Teams and Race Relations in the Classroom: Four Field Experiments Using Teams–Games–Tournament. *Journal of Educational Psychology*, 70(3), 356–362.
- DeVries, D., Edwards, K. & Wells, E. (1974). *Teams–Games in the Social Studies Classroom: Effects on Academic Achievement, Student Attitudes, Cognitive Beliefs and Classroom Climate* (Informe No. 173). Baltimore: The Johns Hopkins University.
- Díaz–Aguado, M. J. (1996). Aprendizaje cooperativo y experiencias de responsabilidad. En *Programas de educación para la tolerancia y prevención de la violencia en los jóvenes. Vol. 1. Fundamentación Psicopedagógica*. Madrid: Ministerio de Trabajo y Asuntos Sociales, Instituto de la Juventud.
- Díaz–Aguado, M. J. y Andres, T. (1999) Aprendizaje cooperativo y educación intercultural. Investigación–acción en centros de primaria. *Psicología Educativa*, 5(2), 141–200.
- Dickersin, K. (1994). Sobre la existencia y los factores de riesgo del sesgo de publicación. *Boletín de la Oficina Sanitaria Panamericana*, 116, 435–446.
- Díez, E. (2011). Modelos socioconstructivistas y colaborativos en el uso de las TIC en la formación inicial del profesorado. *Revista Educación*, 358, 175–196.
- Dillenbourg, P. (2003). *Collaborative Learning. Cognitive and Computational Approaches*. Hungría: Earli.
- Dioso–Henson, L. (2012). The effect of Reciprocal Peer Tutoring and non–Reciprocal Peer Tutoring on the performance of students in college physics. *Research in Education*, 87, 34–49.
- \*Doymuş, K. (2007). Effects of a cooperative learning strategy on teaching and learning phases of matter and one–component phase diagrams. *Journal of Chemical Education*, 84(11), 1857–1860.
- \*Doymuş, K. (2008a). Teaching chemical bonding through jigsaw cooperative learning. *Research in Science & Technological Education*, 26, 47–57.
- Doymuş, K. (2008b). Teaching chemical equilibrium with the Jigsaw Technique. *Research Science Education*, 38, 249–260.
- \*Doymuş, K., Karacop, A. & Simsek, U. (2010). Effects of *Jigsaw* and animation techniques on students' understanding of concepts and subjects in

- electrochemistry. *Educational Technology Research and Development*, 58, 671–691.
- \*Doymuş, K., Şimşek, U. & Karaçöp, A. (2009). The effects of computer animations and cooperative learning methods in micro, macro and symbolic level learning of states of matter. *Eurasian Journal of Educational Research*, 36, 109–128.
- Dufrene, B., Noell, G., Gilbertson, D. & Duhon, G. (2005). Monitoring Implementation of Reciprocal Peer Tutoring: Identifying and Intervening With Students Who Do Not Maintain Accurate Implementation. *School Psychology Review*, 34(1), 74–86.
- Duran, D. (2006). Tutoría entre iguales, la diversidad en positivo. Aula de Innovación Educativa. *Revista Aula de Innovación Educativa*, 153–154.
- Duran, D. (2012). Estructurar la interacción a través de métodos y técnicas. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.139–166). Madrid: Alianza editorial.
- Duran, D. y Monereo, C. (2012). *Entramado. Métodos de aprendizaje cooperativo y colaborativo*. Barcelona: Horosori Editorial.
- Duran, D., Torró, J. y Vila, J. (2003). *Tutoría ente iguales*. Barcelona: ICE de la UAB.
- Duran, D. y Vidal, C. (2004). *Tutoría entre iguales: de la teoría a la práctica. Un método de aprendizaje cooperativo para la diversidad en secundaria*. España: Grao.
- Durlak, J. & Lipsey, M. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology*, 19(3), 291–332.
- Echeita, G. (1995). El aprendizaje cooperativo. Un análisis psicosocial de sus ventajas respecto a otras estructuras de aprendizaje. En P. Fernández y A. Melero (Comps). *La interacción social en contextos educativos* (pp.167-189). Madrid: Siglo XXI.
- Echeita, G. (2012). El aprendizaje cooperativo al servicio de una educación de calidad. Cooperar para aprender y aprender para cooperar. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.21–45). Madrid: Alianza editorial.
- Echeita, G. y Martín, E. (1990). Interacción social y aprendizaje. En C. Coll, J. Palacios y A. Marchesi (Comp.) *Desarrollo psicológico y educación. Psicología de la educación. Vol. III* (pp.49-67). Madrid: Alianza Psicología.

- Education Group on Guidelines on Evaluation (1999). Guidelines for evaluating papers on educational interventions. *BMJ*, 318, 1265–1267.
- Egger, M., Smith, G., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 315, 629-634.
- \*Emley, W. (1986). The effectiveness of cooperative learning versus individualized instruction in a college level remedial mathematics course, with relation to attitudes toward mathematics and Myers–Briggs personality type (Tesis doctoral). The University of Maryland, USA.
- Enesco, I. y Del Olmo, C. (1992). *El trabajo en equipo en primaria, aprendiendo con iguales*. Madrid: Alhambra Longman.
- Escudero, J. (2012). La colaboración docente, una manera de aprender juntos sobre el trabajo cooperativo en el alumnado. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.269–288). Madrid: Alianza editorial.
- Evidence Informed Policy and Practice in Education in Europe (2011). *The nature of evidence*. Recuperado de <http://www.eippee.eu/cms/Default.aspx?tabid=3179>.
- Eysenck, H. J. (1994). Systematic reviews: Meta-analysis and its problems. *British Medical Journal*, 309, 789–792.
- Fabra, M. L. (1994). *Técnicas de grupo para la cooperación*. Barcelona: CEAC
- Falchikov, N. (2001). *Learning together. Peer tutoring in higher education*. New York: RoutledgeFalmer.
- Farrington, D. (2003). Methodological quality standards for evaluation research. *ANNALS*, 587, 49-68.
- Fantuzzo, J., Dimeff, L. & Fox, S. (1989). Reciprocal Peer Tutoring: A Multimodal Assessment of Effectiveness With College Students. *Teaching of Psychology*, 16(3), 13–135.
- Fantuzzo, J., King, J. & Heller, L. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology*, 84(3), 331–339.
- \*Fantuzzo, J., Riggio, R. E., Connelly, S. & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: a component analysis. *Journal of Educational Psychology*, 81(2), 173–177.
- Faust, J. & Paulson, D. (1998). Active learning in the college classroom. *Journal on Excellence in College Teaching*, 9(2), 3–24.

- Fay, T. & Hardie, K. (2003). Fahrenheit to Celsius: An Exploration in College Algebra. *Mathematics and Computer Education*, 37(2), 207–209.
- \*Feyzioğlu, B., Akçay, H. & Şahin-Pekmez, E. (Agosto, 2007). *Comparison of computer assisted cooperative, competitive and individualistic learning: an example of Turkey*. Congrès international AREF. Actualité de la recherche en education et en formation, Strasbourg, Francia.
- Fernández, I., Guisasola, G., Garmendia, M., Alkorta, I. y Madinabeitia, A. (2013). ¿Puede la formación docente tener efectos globales en la universidad? Desarrollo docente, metodologías activas y currículum híbrido. *Infancia y aprendizaje*, 36(3), 387–400.
- Fernández, R. (2004). Hacia un nuevo paradigma educativo: análisis de estrategias de aprendizaje colaborativo en la formación inicial de maestros en nuevas tecnologías aplicadas a la educación en la escuela universitaria de Magisterio de Toledo. *Revista Latinoamericana de Tecnología Educativa*, 3(1), 195–212.
- Ferreiro, R. (2006). *Estrategias didácticas del aprendizaje cooperativo. El constructivismo social: una forma de enseñar y de aprender*. México: Trillas.
- Ferreiro, R. (2010). *Estrategias didácticas del aprendizaje cooperativo. Método ELI*. México: Trillas.
- Foyle, H. & Shafto, M. (1995). Teamwork in the real world. En H. Foyle (Edit.) *Interactive Learning in the Higher Education Classroom. Cooperative, Collaborative, and Active Learning Strategies* (pp.11–17). Washington: The NEA Professional Library Higher Education Series.
- Fraile, A. (2008). El aprendizaje cooperativo como metodología para el desarrollo de los ECTS: una experiencia de formación del profesorado en educación física. *Revista Fuentes*, 8, 22–35.
- Fritz, C., Morris, P. & Richler, J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: general*, 141(1), 2-18.
- \*Ganter, S. (1994). The importance of empirical evaluations of mathematics programs: a case from the calculus reform movement. *Focus on Learning Problems in Mathematics*, 16(2), 1–19.
- García, A. y Suárez, C. (2011). Interacción virtual y aprendizaje cooperativo. Un estudio cualitativo. *Revista de Educación*, 354, 473–498.
- García Hoz, V. (1972). *Técnicas de trabajo cooperativo en la enseñanza universitaria*. Madrid: Gráficas Cóndor.

- Garton, A. (1994). *Interacción social y desarrollo del lenguaje y la cognición*. Barcelona: Paidós.
- Gavilán, B. P. (1997). El aprendizaje cooperativo: Desde las matemáticas también es posible educar en valores. *Revista de Didáctica de las Matemáticas*, 13, 81–94.
- Gavilán, P. (2004). *Álgebra en secundaria. Trabajo cooperativo en matemáticas*. Madrid: Ministerio de Educación y Ciencia, Narcea.
- Gavilán, P. (2004). *Materiales 12–16 para educación secundaria: Álgebra en secundaria. Trabajo cooperativo en matemáticas*. Madrid: Ministerio de Educación y Ciencia, Narcea.
- Gavilán, P. (2009). Aprendizaje cooperativo. Papel del conflicto sociocognitivo en el desarrollo intelectual. Consecuencias pedagógicas. *Revista Española de Pedagogía*, 242, 131–148.
- Gavilán, P. y Alario, R. (2010). *Aprendizaje Cooperativo. Una metodología con futuro. Principios y aplicaciones*. Madrid: Editorial CCS.
- Geok-Chin, I., Kim-Eng, C. & Sharan, S. (2007). Group Investigation Effects on Achievemem, Motivation, and Perceptions of Students in Singapore. *The Journal of Educational Research*, 100(3), 142–154.
- Geok-Chin, I., Sharan, S. & Kim-Eng, C. (2005). Student's Perceptions of Learning Geography through Group Investigation in Singapore. *International Research in Geographical and Environmental Education*, 14(4), 261–276.
- \*George, P. (1994). The Effectiveness of Cooperative Learning Strategies in Multicultural University Classrooms. *Journal on Excellence in College Teaching*, 5(1), 21–30.
- Gerard, H., Wilhelmy, R. & Conolley, E. (1968). Conformity and group size. *Journal of Personality and Social Psychology*, 8(1), 79–82.
- Gessa, A. (2011). La coevaluación como metodología complementaria de la evaluación del aprendizaje. Análisis y reflexión en las aulas universitarias. *Revista de Educación*, 354, 749–764.
- \*Ghaith, G. & El-malak, M. (2004). Effect of Jigsaw II on literal and higher order EFL reading comprehension. *Educational Research and Evaluation*, 10(2), 105–115.
- Gil, C., Baños, R., Montoya, M., Herrada, R. y Montoya, G. (2013). Aprendizaje Cooperativo y Portafolio Digital de Grupo: Desarrollo de Competencias. *Enseñanza y Aprendizaje de Ingeniería de Computadores*, 3, 23–37.

- Gil, F., Rico, R. y Sánchez-Manzanares, M. (2008). Eficacia de equipos de trabajo. *Papeles del Psicólogo*, 29(1), 25–31.
- Gillespie, A. & Richardson, B. (2011). Exchanging social positions: Enhancing perspective taking within a cooperative problem solving task. *European of Social Psychology*, 41, 608–616.
- Ginsburg-Block, M., Rohrbeck, C. & Fantuzzo, J. (2006). A meta-analytic review of social, self-concept and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, 98(49), 732–749.
- \*Giraud, G. (1997). Cooperative Learning and Statistics Instruction. *Journal of Statistics Education*, 5(3), 1–11.
- Glass, G. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher*, 5, 3–8.
- Glass, G. (1977). Integrating findings: the meta-analysis of research. *Review of Research in Education*, 5, 351–379.
- Glass, G., McGaw, B. & Smith, M. (1981). *Meta-analysis in social research*. Beverly-Hills, CA: Sage.
- Goikoetxea, E. y Pascual, G. (2002). Aprendizaje cooperativo: bases teóricas y hallazgos empíricos que explican su eficacia. *Educación XXI: Revista de la Facultad de Educación*, 5, 199–226.
- Goldschmid, B. & Goldschmid, M. (1976). Peer teaching in higher education: a review. *Higher Education*, 5, 9–33.
- Gómez, C., Palomares, R. y Pino, J. (2010). La utilización de herramientas colaborativas 2.0 en el ámbito de la documentación publicitaria. *Didáctica, Innovación y Multimedia*, 18, 1–8.
- \*Gömleksiz, M. (2007). Effectiveness of cooperative learning (Jigsaw II) method in teaching English as a foreign language to engineering students (Case of Firat University, Turkey). *European Journal of Engineering Education*, 32(5), 613–625.
- González, E. y Zariquiey, F. (2012). Las TIC y el aprendizaje cooperativo. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.207–238). Madrid: Alianza editorial.



- Goodman, K. & Pascarella, E. (2006). First-year seminars increase persistence and retention: A summary of the evidence from How College Affects Students. *Peer Review*, 8(3), 26–28.
- Goodsell, A., Maher, M. & Tinto, V. (1992). *Collaborative learning: a sourcebook for higher education*. Guide-Non Classroom Use. Recuperado de <http://eric.ed.gov/?id=ED357705>
- Graham, S. & Perin, D. (2007). A Meta-Analysis of Writing Instruction for Adolescent Students. *Journal of Educational Psychology*, 99(3), 445-476.
- Graves, T. (1991). The controversy over group rewards in cooperative classrooms. *Educational Leadership*, 77–79.
- Greenhouse, J. & Iyengar, S. (2004). Sensitivity analysis and diagnostics. En H. Cooper, L. Hedges & J. Valentine (Eds.) *The Handbook of Research Synthesis* (2ª Ed.) (pp.417–433). New York: Russell Sage Foundation.
- Güneysu, S. & Tekmen, B. (2010). Implementing an alternative cooperative learning method. *Procedia Social and Behavioral Sciences*, 2, 5670–5674.
- Gupta, A., Mattarelli, E., Seshasai, S. & Broschak, J. (2009). Use of collaborative technologies and knowledge sharing in co-located and distributed teams: Towards the 24-h knowledge factory. *Journal of Strategic Information Systems*, 18, 147–161.
- Gutiérrez, I. (2009). El maestro de la “Experiencia Somosaguas”. *Tendencias Pedagógicas*, 14, 181–190.
- Gutiérrez, P., Yuste, S., Cubo, S. y Fustes, M. (2011). Buenas prácticas en el desarrollo de trabajo colaborativo en materias TIC aplicadas a la educación. *Revista de Currículum y Formación del Profesorado*, 15(1), 179–194.
- Haas, M. (2005). Teaching Methods for Secondary Algebra: A Meta-Analysis of Findings. *NASSP Bulletin*, 89, 24-46
- Hagelgans, N., Reynolds, B., Schwingendorf, K., Vidakovic, D., Shahin, M. & Wimbish, G. (1995). *A practical guide to Cooperative Learning in Collegiate Mathematics*. Washington: The Mathematical Association of America.
- Hanshaw, L.G. (1982). Test anxiety, self-concept, and the test performance of students paired for testing and the same students working alone. *Science Education*, 66, 15–24.

- \*Harding, R. & Fletcher, R. (Noviembre, 1994). Effectiveness of variations in collaborative cooperative learning in RDS mathematics classes. Paper presented at the Annual Meeting of the Tennessee Academy of Science, Nashville, USA.
- Hargreaves, D. (2011). Teaching as a research-based profession: possibilities and prospects. En M. Hammersley (Ed.) *Educational Research and Evidence-based Practice* (pp.3–17). United Kingdom: The Open University.
- Hart Research Associates (2008). *How should colleges assess and improve student learning? Employers' views on the accountability challenge: A survey of employers conducted on behalf of the Association of American Colleges and Universities*. Recuperado de <http://www.aacu.org>
- \*Hassan, M., Fong, S. & Idrus, R. (Octubre, 2011). *Impact of E-Cooperative Learning Modules on Interpersonal Communication Skills. Recent Researches in Education*. 10<sup>th</sup> WSEAS International Conference on Education and Educational Technology (EDU '11). Malaysia.
- Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. New York: Routledge.
- Hattie, J., Rogers, J. & Swaminathan, H. (2014). The role of meta-analysis in Educational Research. En A. Reid, P. Hart & M. Peters (Eds.) *A companion to Research in Education* (pp.197-207). USA: Springer Science+Business Media Dordrecht (e-book) Recuperado de [http://link.springer.com/chapter/10.1007/978-94-007-6809-3\\_26](http://link.springer.com/chapter/10.1007/978-94-007-6809-3_26)
- Hedges, L. & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.
- Hedges, L. & Pigott, T. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9(4), 426–445.
- Hernández, R., Fernández, C. y Baptista, P. (2010). *Metodología de la investigación* (5<sup>a</sup> Ed.). México: McGraw-Hill.
- Hernando, A., Aguaded, J. y Tirado, R. (2011). Aprendizaje cooperativo on-line a través del campus andaluz virtual. Análisis de las interacciones. *Enseñanza & Teaching*, 29(1), 135–158.
- Hertz-Lazarowitz, R. (1985). Internal dynamics of cooperative learning groups. En R. Slavin, S. Sharon, S. Kagan, R. Lazarowitz, C. Webb & R. Schmuck. *Learning to cooperate, cooperating to learn* (pp.97–102). Plenum Press: New York.

- Hertz–Lazarowitz, R., Benveniste, V. & Miller, N. (1995). An Overview of the Theoretical Anatomy of Cooperation. En R. Hertz–Lazarowitz & N. Miller (Edit.). *Interaction in Cooperative Groups. The Theoretical Anatomy of Group Learning* (pp.1-14). United States of America: Cambridge University Press.
- Higgins, J. (2008). Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *Journal of Epidemiology*, 37, 1158–1160.
- Higgins, J. & Green S. (2011) (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration, 2011. Recuperado de [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
- Higgins, J. & Thompson, S. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 15(21), 1539–1558.
- Higgins, J., Thompson, S., Deeks, J. & Altman, D. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557 – 560.
- Hogan, D. & Tuge, J. (1999). Implications of Vygotsky’s theory for peer learning. En A. O’Donnell & A. King (Eds.) *Cognitive perspectives on peer learning* (pp.3–38). New York: Routledge.
- Hogan, T. P. (2004). *Pruebas psicológicas. Una introducción práctica*. México: Manual Moderno.
- Holliday, D. (2002). *Jigsaw IV: Using Student/Teacher Concerns To Improve Jigsaw III*. Guides–Classroom–Teacher. Recuperado de <http://files.eric.ed.gov/fulltext/ED465687.pdf>.
- Hooper, S. (1992). Cooperative Learning and Computer–Based Instruction. *Educational Technology Research and Development*, 40(3), 21–38.
- Hopewell, S., Loudon, K., Clarke, M.J., Oxman, A.D. & Dickersin, K. (2009). Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews 1*. DOI: 10.1002/14651858.MR000006.pub3.
- Huang, T., Huang, Y. & Yu, F. (2011). Cooperative Weblog Learning in Higher Education: Its Facilitating Effects on Social Interaction, Time Lag, and Cognitive Load. *Educational Technology & Society*, 14(1), 95–106.
- Huang, Y.–M., Huang, T.–C. & Hsieh, M.–Y. (2008). Using annotation services in a ubiquitous Jigsaw cooperative learning environment. *Educational Technology & Society*, 11(2), 3–15.

- Huber, G. (2008). Aprendizaje activo y metodologías educativas. *Revista de Educación*, número extraordinario, 59–81.
- Huedo, T. (2006). Índices de tamaño del efecto para meta-analizar diseños pretest–posttest con dos grupos: Un estudio de simulación Monte Carlo. [Effect Size indices for meta-analysis with two-groups pretest–posttest designs: a Monte Carlo simulation study]. Unpublished Doctoral Dissertation. Universidad Nacional de Educación a Distancia (UNED), Spain.
- Huedo–Medina, T. y Johnson, B. (2010). *Modelos estadísticos en meta-análisis*. España: Netbiblo.
- Huedo–Medina, T., Sánchez–Meca, J., Marin, F. & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I<sup>2</sup> index? *CHIPS document*. Paper 19. University of Connecticut. Recuperado de [http://digitalcommons.uconn.edu/cgi/viewcontent.cgi?article=1019&context=chip\\_docs](http://digitalcommons.uconn.edu/cgi/viewcontent.cgi?article=1019&context=chip_docs)
- Huhtala, J. & Coughlin, E. (1991). Group Investigation, Democracy, and the Middle East. *English Journal*, 80(5), 47–52.
- \*Hui–Chuan, L. (2005). *Effects of cooperative learning on motivation, learning strategy utilization, and grammar achievement of English language learners in Taiwan* (Tesis doctoral). University of New Orleans, USA.
- Hunter, J. & Schmidt, F. (1996). Cumulative research knowledge and social policy formulation: the critical role of meta-analysis. *Psychology Public Policy and Law*, 2(2), 324–347.
- \*Hsiung, C. (2011). Identification of dysfunctional cooperative learning teams using Taguchi Quality Indexes. *Educational Technology & Society*, 14(3), 152–162.
- \*Hsiung, C. (2012). The Effectiveness of Cooperative Learning. *Journal of Engineering Education*, 101(1), 119–137.
- \*Hwong, N., Caswell, A., Johnson, D. & Johnson, R. (1993). Effects of cooperative and individualistic learning on prospective elementary teachers' music achievement and attitudes. *Journal of Social Psychology*, 133(1), 53–64.
- Iborra, M. y Dasí, A. (2009). ¿Cómo aprender a trabajar en equipos integrados? El papel del aprendizaje cooperativo en la formación de directivos. *Revista D'Innovació educativa*, 3, 29–38.

- Institute of Education Sciences (2010). *The What Works Clearinghouse WWC Evidence review protocol for character education interventions*. Recuperado de <http://ies.ed.gov/ncee/wwc/document.aspx?sid=23>
- \*Iwasiw, C. L. & Goldenberg, D. (1993). Peer teaching among nursing students in the clinical area: effects on student learning. *Journal of Advanced Nursing*, 18, 659–68.
- \*Jalilifar, A. (2010). The effect of cooperative learning techniques on college students' reading comprehension. *System*, 38, 96–108.
- Jansoon, N., Somsook, E. & Coll, R. (2008). Thai Undergraduate Chemistry Practical Learning Experiences Using the Jigsaw IV Method. *Journal of Science and Mathematics Education in Southeast Asia*, 31(2), 178–200.
- Jessen–Marshall, A. & Lescinsky, H. (2011). Origins. Team Teaching in the Sciences. En K. Plank. *Team Teaching. Across the Disciplines, Across the Academy* (pp.13–36). Virginia: Stylus – The National Teaching and Learning Forum.
- Jiménez, R. (2012). Educación inclusiva y formación inicial del profesorado: evaluación de una innovación didáctica basada en la producción cinematográfica desde la perspectiva del alumnado. *Revista Educación*, 359, 232–259.
- Jiménez-Contreras, E. (2000). Los métodos bibliométricos. Estado de la cuestión y aplicaciones. *Cuadernos de documentación multimedia*, 10, 757–771.
- \*Jimison, L. (1990). A study to investigate the effect of cooperative group learning on selected cognitive and affective outcomes for preservice elementary teachers in a mathematics methods class. (Tesis doctoral). Oklahoma State University, USA.
- Johns Hopkins University (2014). *Best Evidence Encyclopedia (BEE): Standards for Research–Proven Programs and Frequently Asked Questions*. Recuperado de <http://www.bestevidence.org/resources/general/standards.htm>
- Johnson, D. & Johnson, R. (1985). The internal dynamics of cooperative learning groups. En R. Slavin, S. Sharon, S. Kagan, R. Lazarowitz, C. Webb & R. Schmuck (Ed.). *Learning to cooperate, cooperating to learn* (pp.103–124). Plenum Press: New York.
- Johnson, D. & Johnson, R. (1989). *Cooperation and competition. Theory and research* (2ª Ed.). Minesota: Interaction Book Company.
- Johnson, D. & Johnson, R. (1994). Learning together. En S. Sharan (Ed.) *Handbook of cooperative learning methods* (pp.51–65). USA: Praeger.

- Johnson, D. & Johnson, R. (1999a). Learning Together. En S. Sharan (Edit.), *Handbook of Cooperative Learning Methods* (pp.51–65). Westport–Connecticut: Praeger.
- Johnson, D. & Johnson, R. (1999b). *Learning Together and Alone. Cooperative, Competitive, and Individualistic Learning* (5ª Ed.). Boston: Allyn and Bacon.
- Johnson, D. y Johnson, R. (1999c). *Aprender juntos y solos. Aprendizaje cooperativo, competitivo e individualista*. Argentina: AIQUE
- Johnson, D. & Johnson, R. (1999d). Making cooperative learning work. *Theory into practice*, 38(2), 67–73.
- Johnson, D. & Johnson, R. (1999e). An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. *Educational Researcher*, 38, 365–379.
- Johnson, D. & Johnson, R. (2002a). Social interdependence theory and university instruction: Theory into practice. *Swiss Journal of Psychology*, 61, 119–129.
- Johnson, D. & Johnson, R. (2002b). Cooperative learning methods: a meta-analysis. *Journal of Research in Education*, 12(1), 5–24.
- Johnson, D. & Johnson, R. (2009a). Energizing Learning: The Instructional Power of Conflict. *Educational Researcher*, 38, 37–51. Recuperado de <http://edr.sagepub.com/content/38/1/37>.
- Johnson, D. & Johnson, R. (2009b). An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. *Educational Researcher*, 38: 365–379. Recuperado de: <http://edr.sagepub.com/cgi/content/abstract/38/5/365>.
- Johnson, D. & Johnson, R. (2014a). *Joining Together: Group Theory and Group Skills* (11ª Ed.). Englewood Cliffs, Nj: Prentice–Hall.
- Johnson, D. y Johnson, R. (2014b). Taller teórico–práctico de aprendizaje cooperativo. Taller organizado por la Fundación CVE Educar. Madrid.
- Johnson, D., Johnson, R. & Holubec, E. (1992). *Advanced cooperative learning*. Minnesota: International Book Company.
- Johnson, D., Johnson, R. & Holubec, E. (1999). *Los nuevos círculos del aprendizaje. La cooperación en el aula y la escuela*. Argentina: AIQUE
- Johnson, D., Johnson, R. & Smith, K. (1991). *Cooperative Learning: Increasing College Faculty Instructional Productivity*. ASHE–ERIC Higher Education (Report No 4). Washington, D.C.: The George Washington University, Graduate School of Education and Human Development.

- Johnson, D., Johnson, R. y Smith, K. (1997). *El Aprendizaje Cooperativo regresa a la Universidad: ¿qué evidencia existe de que funciona?* Recuperado de [http://www.javeriana.edu.co/prin/sites/default/files/Johnson\\_Aprendizaje\\_cooperativo\\_en\\_la\\_universidad.pdf](http://www.javeriana.edu.co/prin/sites/default/files/Johnson_Aprendizaje_cooperativo_en_la_universidad.pdf)
- Johnson, D., Johnson, R. & Smith, K. (1998). Cooperative Learning Returns To College: What Evidence Is There That It Works? *Change*, 27–35.
- Johnson, D., Johnson, R. & Smith, K. (2007). The State of Cooperative Learning in Postsecondary and Professional Settings. *Educational Psychology Review* 19, 15–29.
- Johnson, D. W., Johnson, R., Stanne, M. & Garibaldi, A. (1990). The impact of leader and member group processing on achievement in cooperative groups. *Journal of Social Psychology*, 130, 507–516.
- Johnson, D., Johnson, R. & Stanne, M. (2000). *Cooperative learning methods: A meta-analysis*. Cooperative Learning Center, University of Minnesota. Recuperado de <http://www.co-operation.org/pages/cl-methods.html>.
- Johnson, D., Johnson, R. & Tjosvold, D. (Febrero, 2008). *Constructive controversy: the value of intellectual conflict*. Trabajo presentado en la IACM 21<sup>st</sup> Annual Conference Paper. Recuperado de <http://ssrn.com/abstract=1298645>
- Johnson, D., Maruyama, G., Johnson, R., Nelson, D. & Skon, L. (1981). Effects of Cooperative, Competitive, and Individualistic Goal Structures on Achievement: A Meta-Analysis. *Psychological Bulletin*, 981(89), 47–62.
- Johnston, S. (2009). A crisis of clarity. En J. Cooper, P. Robinson & D. Ball *Small Group Instruction in Higher Education. Lessons from the Past, Visions of the Future*. (2<sup>a</sup> Ed.). (pp.349–356). Sitllwater: New Forums.
- Jones, K. & Jones, J. (2008). Making Cooperative Learning Work in the College Classroom: An Application of the ‘Five Pillars’ of Cooperative Learning to Post-Secondary Instruction. *The Journal of Effective Teaching*, 8(2), 61–76.
- Jüni, P., Witschi, A., Bloch, R. & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*, 282, 1054-1060.
- Kacer, B., Rocklin, T. & Weinhloltz, D. (1992). Individual versus small group instruction of computer applications: A quantitative and qualitative comparison. *Journal of Computing in Teacher Education*, 9(1), 6–12.

- Kagan, S. (1985). Dimensions of cooperative classroom structures. En R. Slavin, S. Sharon, S. Kagan, R. Lazarowitz, C. Webb & R. Schmuck (Eds.) *Learning to cooperate, cooperating to learn* (pp.67–95). New York: Plenum Press.
- Kagan, S. (1990). The structural approach to cooperative learning. *Educational Leadership*, 12–14.
- Kagan, S. (1994). *Cooperative learning*. San Clemente, CA: Kagan Publishing.
- Kagan, S. & Kagan, M. (2009). *Kagan cooperative learning*. San Clemente, CA: Kagan Publishing.
- Kagan, S. & Kagan, M. (2011). Five must-know Kagan structure for higher education. En J. Cooper & P. Robinson *Small group learning in Higher Education* (pp.109–113). United States of America: New Forums Press.
- Kalain, S. & Kasim, R. (2014). A meta-analytic review of studies of the effectiveness of small-group learning methods on statistics achievement. *Journal of Statistics Education*, 22(1), 1-20.
- Kapp, E. (2009). Improving Student Teamwork in a Collaborative Project-Based Course. *College Teaching*, 57(3), 139–143.
- \*Karababa, C. (2009). Effects of cooperative learning on prospective teachers' achievement and social interactions. *Hacettepe Üniversitesi Eğitimi Fakültesi Dergisi (H. U. Journal of Education)*, 36, 32–40.
- Karacop, A. & Doymuş, K. (2013). Effects of Jigsaw Cooperative Learning and Animation Techniques on Students' Understanding of Chemical Bonding and Their Conceptions of the Particulate Nature of Matter. *Journal of Science Education and Technology*, 22(2), 186–203.
- Kavale, K.A. & Glass, G.V. (1981). Meta-analysis and the integration of research in special education. *Journal of Learning Disabilities*, 14, 531–538.
- \*Keeler, C. & Steinhurst, R. (1994). Cooperative learning in statistics. *Teaching Statistics*, 16(3), 81–84.
- Kerlinger, F. (2002). *Investigación del Comportamiento* (3ª Ed.). México: McGraw-Hill.
- \*Kerr, D. & Murthy, U. (1994). Group decision systems and cooperative learning in auditing: an experimental investigation. *Journal of Information Systems*, 8(2), 85–96.



- \*Kiliç, D. (2008). The effects of the Jigsaw Technique on learning the concepts of the principles and methods of teaching. *World Applied Sciences Journal* 4 (supple 1), 109–114.
- \*Klein, J. & Pridemore, D. (1992). Effects of Cooperative Learning and Need for Affiliation on Performance, Time on Task, and Satisfaction. *Educational Technology Research and Development*, 40(4), 39–47.
- \*Klein, J. & Schnackenberg, H. (2000). Effects of informal cooperative learning and the affiliation motive on achievement, attitude, and student interactions. *Contemporary Educational Psychology*, 25, 332–341.
- Knight, G. & Morton, E. (1990). Cooperative Learning and Achievement: Methods for Assessing Causal Mechanisms. En S. Sharan (Eds.) *Cooperative Learning: Theory and Research* (pp.1–22). Westport, Connecticut: Praeger.
- \*Koç, Y., Doymuş, K., Karaçöp, A. & Şimşek, U. (2010). The Effects of Two Cooperative Learning Strategies on the Teaching and Learning of the Topics of Chemical Kinetics. *Journal of Turkish Science Education*, 7(2), 52–65.
- Konstantopoulos, S. & Hedges, L. (2004). Analyzing effect sizes: fixed-effects models. En H. Cooper, L. Hedges & J. Valentine (Eds.) *The Handbook of Research Synthesis* (2<sup>a</sup> Ed.) (pp.279–293). New York: Russell Sage Foundation.
- Kozlowski, S. & Ilgen, D. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, 7(3), 77–124.
- \*Krause, U., Stark, R. & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, 19, 158–170.
- Kulik, C. & Kulik, J. (1986). Effectiveness of computer-based education in colleges. *AEDS Journal*, 19, 81–108.
- Kulik, C. & Kulik, J. (1989). Meta-analysis in education. *International Journal of Educational Research*, 13, 221–340.
- Kulik, C. & Kulik, J. (1991). Effectiveness of computer-based instruction: an updated analysis. *Computer in Human Behavior*, 7, 75–94.
- Kulik, C. Kulik, J. & Cohen, P. (1980). Effectiveness of computer-based college teaching: a meta-analysis of findings. *Review of Educational Research*, 50, 525–544.
- Kulik, C., Schwalb, B. & Kulik, J. (1982). Programmed Instruction in Secondary Education: A Meta-Analysis of Evaluation Findings. *Journal of Educational Research*, 1(75), 133–138.

- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E. & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review*, 10, 133–149.
- Laboratorio de innovación educativa (Lab) (2009). Qué, por qué, para qué, cómo. *Aprendizaje cooperativo. Propuesta para la implantación de una estructura de cooperación en el aula*. Recuperado de [http://crei.centros.educa.jcyl.es/sitio/upload/LAB\\_DOCUMENTACION\\_APRENDIZAJE\\_COOPERATIVO.pdf](http://crei.centros.educa.jcyl.es/sitio/upload/LAB_DOCUMENTACION_APRENDIZAJE_COOPERATIVO.pdf)
- Lai, C.-Y. & Wu, C.-C. (2006). Using handhelds in a *Jigsaw* cooperative learning environment. *Journal of Computer Assisted Learning*, 22, 284–297.
- \*Lambiotte, J., Dansereau, D., Rocklin, T., Fletcher, B., Hythecker, V., Larson, C. & O'Donnell, A. (1987). Cooperative learning and test taking: transfer of skills. *Contemporary Educational Psychology*, 12, 52–61.
- \*Lambiotte, J., Skaggs, L. & Dansereau, D. (1993). Learning from Lectures: Effects of Knowledge Maps and Cooperative Review Strategies. *Applied Cognitive Psychology*, 7, 483–497.
- Landis, J. & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–74.
- León del Barco, B. (2006). Elementos mediadores en la eficacia del aprendizaje cooperativo: entrenamiento previo en habilidades sociales y dinámica de grupos. *Anales de Psicología*, 22(1), 105–112.
- León del Barco, B. y Latas, C. (2007). La formación en técnicas de aprendizaje cooperativo del profesor universitario en el contexto de la convergencia europea. *Revista de Psicodidáctica*, 12(2), 269–278.
- León del Barco, B., Felipe, E., Iglesias, D. y Latas, C. (2011). El aprendizaje cooperativo en la formación inicial del profesorado de Educación Secundaria. *Revista de Educación*, 354, 715–729.
- León del Barco, B., Felipe, E., Iglesias, D. y Marugán, M. (2014). Determinantes en la eficacia del aprendizaje cooperativo. Una experiencia en el EEES. *Revista de Investigación Educativa*, 32(2), 411–424.
- León del Barco, B., Gozalo, M. y Vicente, F. (2004). Factores mediadores en el aprendizaje cooperativo: los estilos de conducta interpersonal. *Apuntes de Psicología*, 22(1), 61–74.

- León del Barco, B., Gozalo, M., Felipe, E., Gómez, T. y Latas, C. (2005). *Técnicas de aprendizaje cooperativo en contextos educativos*. Badajoz: Editorial @becedario.
- \*Leveson, L. (1999). Small group work in accounting education: an evaluation of a programme for first-year students. *Higher Education Research & Development*, 18(3), 361–377.
- Levine, T., Asada, K. & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analyses: evidence and implications of a publication bias implications of a publication bias against nonsignificant findings. *Communication Monographs*, 76(3), 286–302.
- Liao, M. & Worth, S. (2011). Lessons Learned by a Philosopher and a Biologist in Team Teaching a First-Year Seminar on “Disease and Culture: Why You Are a Walking Petri Dish”. En K. Plank. *Team Teaching. Across the Disciplines, Across the Academy* (pp.37–54). Virginia: Stylus – The National Teaching and Learning Forum.
- \*Lin, Z. (2012). Comparison of technology-based cooperative learning with technology-based individual learning in enhancing fundamental nursing proficiency. *Nurse Education Today*, 33, 546–551.
- Lincoln, Y. & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage publications.
- Lipsey, M. (1994). Identifying potentially interesting variables and analysis opportunities. En H. Cooper, L. Hedges & J. Valentine (Eds.) *The Handbook of Research Synthesis* (2ª Ed.) (pp.147–158). New York: Russell Sage Foundation.
- Lipsey, M.W. & Wilson, D.B. (1993). The efficacy of psychological, educational and behavioural treatment: Confirmation from meta-analysis. *American Psychologist*, 48(12), 1181–1209.
- Lipsey, M.W. & Wilson, D.B. (2001). *Practical meta-analysis. Applied Social Research Methods Series*, 45. Thousand Oaks, CA: Sage.
- Liu, C.-C., Tao, S.-Y., Nee, J.-N., Liu, B.-J., Chen, G.-D, Hsu, C.-C., Horng, J.-T. (Noviembre, 2005). *ing Activity Awareness for Teams-Games-Tournaments with GSM Network*. Conferencia: Proceedings IEEE International Workshop on Wireless and Mobile Technologies in Education. Tokushima, Japón.
- Llopis, E. (1999). Culturas musicales y aprendizaje cooperativo, pop, rock, heavy y bakalao en la clase de música. *Cuadernos de Pedagogía*, 279, 27–30.

- Lo, H. (2013). Design of Online Report Writing Based on Constructive and Cooperative Learning for a Course on Traditional General Physics Experiments. *Educational Technology & Society*, 16(1), 380–391.
- Lobato, C. (1997). Hacia una comprensión del aprendizaje cooperativo. *Revista de Psicodidáctica*, 4, 59–76.
- Lobato, C. (1998). *Aprendizaje cooperativo en secundaria*. Bilbao: Servicio Editorial de Universidad del País Vasco.
- Lobato, C., Apodaca, P., Baradian, M., San José, M., Sancho, J. & Zubimendi, J. (2010). Development of the competences of teamwork through cooperative learning at the university. *International Journal of Information and Operations Management Education*, 3(3), 224–240.
- López, D. y Álvarez, I. (2011). Promover la regulación del comportamiento en tareas de aprendizaje cooperativo en línea a través de la evaluación. *Ried*, 14(1), 161–183.
- López, F. (2005). *Metodología participativa en la enseñanza universitaria*. Madrid: Narcea.
- López, G., Ojea, M. y Fernández, C. (2000). Inclusión educativa de estudiantes con necesidades educativas a través del aprendizaje cooperativo. *Aula de Innovación Educativa*, 90, 36–38.
- López, L. (2012). La alfabetización mediática a debate. Planteamiento didáctico de una dinámica de grupo en el ámbito universitario. *Icono 14*, 10(3), 85–99.
- Lou, Y., Abrami, P. & D' Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71(3), 449–521.
- Lou, Y., Abrami, P. & Spence, J. (2000). Effects of within-class grouping on student achievement: an exploratory model. *The Journal of Educational Research*, 94(2), 101–112.
- Lou, Y., Abrami, P., Spence, J., Poulsen, C., Chambers & D'Apollonia, S. (1996). Within-class grouping: a meta-analysis. *Review of Educational Research*, 66(4), 423–458.
- \*Lovelace, T. L. & McKnight, C. K. (1980). The effects of reading instruction on calculus students' problem solving. *Journal of Reading*, 23, 305–8.
- Lynch, B. L. (1984). Cooperative learning in interdisciplinary education for the allied health professions. *Journal of Allied Health*, 13(2), 83–93.

- MacPherson, R; Jones, A; Whitehouse, C. & O'Neill, P. (2001). Small group learning in the final year of a medical degree: a quantitative and qualitative evaluation. *Medical Teacher*, 23(5), 494–502.
- \*Maden, S. (2010). The effect of Jigsaw IV on the achievement of course of language teaching methods and techniques. *Educational Research and Reviews*, 5(12), 62–71.
- Maier, M., McGoldrick, K. & Simkins, S. (2010). Implementing Cooperative Learning in Introductory Economics Courses. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.157–180). Virginia: Stylus – The National Teaching and Learning Forum.
- Male, M. (1994). Cooperative learning and computer. En S. Sharan (ed) *Handbook of cooperative learning methods* (pp.267–280). USA: Praeger.
- Maloof, J. & White, V. (2005). Team study training in the college biology laboratory. *Journal of Biological Education*, 39(3), 120–124.
- Mansour, N., EL–Deghaidy, H., Alshamrani, S. & Aldahmash, A. (2014 en prensa). Rethinking the Theory and Practice of Continuing Professional Development: Science Teachers' Perspectives. *Research Science Education*.
- Marín, F., Sánchez–Meca, J. y López, J.A. (2009). El metaanálisis en el ámbito de las Ciencias de la Salud: una metodología imprescindible para la eficiente acumulación del conocimiento. *Fisioterapia*, 31(3), 107–114.
- Markulis, P. & Strang, D. (2002). Learning cooperatively may not be learning collaborately! *Developments in Business Simulation and Experiential Learning*, 29, 114–120
- Marqués, P. (2011). *Modelos didácticos de trabajo colaborativo a través de las TIC*. I Congreso de Aprendizaje Colaborativo y TICs (CIMAC). Universidad de Salamanca.
- Martí, E. (1991). *Psicología evolutiva. Teorías y ámbitos de investigación*. Barcelona: Anthropos.
- Martí, E. (1992) “¿De qué depende la eficacia del trabajo en grupo?”. *Aula de Innovación Educativa*, 9.
- Martí, E. (1993). Peer interaction in problema solving. A microgenetic analysis of interpsychological mechanisms. En A. Álvarez, P. del Río y J. Wertsch (Eds. De la serie) y C. Coll y N. Mercer (Eds. Del volumen), *Explorations in socio–*

- cultural studies: Vol. 3. Teaching, learning and interaction* (pp.209–216). Madrid: Fundación infancia y aprendizaje.
- Martí, E. (1997). Trabajamos juntos cuando... *Cuadernos de Pedagogía*, 255, 54–58.
- Martí, E. y Solé, I. (1997). Conseguir un trabajo en grupo eficaz. *Cuadernos de Pedagogía*, 255, 59–64.
- Martínez, J., Cibanal, L. y Pérez, M. (2010). *Metodología y aprendizaje en el Espacio Europeo de Educación Superior. De la Teoría a la Práctica*. Alicante: Publicaciones Universidad de Alicante.
- Más, C., Negro, A. y Torrego, J. (2012). Creación de condiciones para el trabajo en equipo en el aula. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.105–138). Madrid: Alianza editorial.
- Massengill, D. (2011). Promoting Professional Student Learning Through Study Groups: A Case Study. *College Teaching*, 59, 85–92.
- Matsui, T., Kakuyama, T. & Onglatco, M. (1987). Effects of Goals and Feedback on Performance in Groups. *Journal of Applied Psychology*, 72(3), 407–415.
- Matthews, R., Cooper, J., Davidson, N. & Hawkes, P. (2009). Building bridges between cooperative and collaborative learning. En J. Cooper, P. Robinson y D. Ball. *Small Group Instruction in Higher Education. Lessons from the Past, Visions of the Future*. (2ª Ed.) (pp.6–17). Sitllwater: New Forums.
- Mauri, T., Clarà, M., Ginesta, A. y Colomina, R. (2014). La contribución al aprendizaje en el lugar de trabajo de los equipos docentes universitarios. Un estudio exploratorio. *Infancia y Aprendizaje*, 36(3), 341–360.
- McCarthy, S. & McMahon, S. (1995). From convention to invention: three approaches to peer interaction during writing. En R. Hertz–Lazarowitz & N. Miller (eds.) *Interaction in cooperative groups. The theoretical anatomy of group learning* (pp.1–14). USA: Cambridge University Press.
- \*McDonald, B., Larson, C., Dansereau, D. & Spurlin, J. (1985). Cooperative dyads: impact on text learning and transfer. *Contemporary Educational Psychology*, 10, 369–377.
- McWhaw, K., Schnackernberg, H., Sclater, J. & Abrami, P. (2003). From co–operation to collaboration. Helping students become collaborative learners. En R. Gillies & A. Ashman (Edit.) *Co–operative Learning. The social and intellectual*

- outcomes of learning in groups* (pp.69–86). Estados Unidos de América: Routledge.
- Medina Rivilla, A. (1979). *Dinámica de grupos como método para una educación de las actitudes cooperativas* (Tesis doctoral). Universidad Complutense de Madrid, España.
- Medina Rivilla, A. (2009). Enfoque pluricultural y sistémico de la innovación educativa: modelo para la mejora de las Instituciones Educativas. En A. Medina Rivilla (Ed.). *Innovación de la educación y de la docencia* (pp.19–45). Madrid: Fundación Universitaria Ramón Areces–UNED.
- Medina Rivilla, A., Sevillano, M<sup>a</sup>.L. y de la Torre, S. (coords.). (2009). *Una universidad para el s.XXI. Espacio Europeo de Enseñanza Superior (EEES). Una mirada transdisciplinar, ecoformadora e intercultural*. Madrid: Universitat.
- Mercer, N. (1997). *La construcción guiada del conocimiento. El habla de profesores y alumnos*. Barcelona: Paidós.
- Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology*, 80(1), 1-14.
- Mesmer–Magnus, J., DeChurch, L. Jimenez–Rodríguez, M., Wildman, J. & Shuffler, M. (2011). A meta–analytic investigation of virtuality and information sharing in teams. *Organizational Behavior and Human Decision Processes*, 115, 214–225.
- Millis, B. (2010a). Sequencing Cooperative Learning Activities in Literature Classes. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.141–156). Virginia: Stylus – The National Teaching and Learning Forum.
- Millis, B. (2010b). Why Faculty Should Adopt Cooperative Learning Approaches. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.1–10) Virginia: Stylus – The National Teaching and Learning Forum.
- Millis, B. & Cottell, P. (1998). *Cooperative Learning for Higher Education Faculty*. Ohio: American Council on Education – Oryx Press.
- Ministerio de Educación y Ciencia (2006). *Propuesta para la renovación de las metodologías educativas en la universidad*. Madrid: Secretaria General Técnica.
- Recuperado de [http://infad.eu/RevistaINFAD/2011/n1/volumen3/INFAD\\_010323\\_489-498.pdf](http://infad.eu/RevistaINFAD/2011/n1/volumen3/INFAD_010323_489-498.pdf)

- Molina, I., Gil, E., Bravo, I. y Herrero, R. (2010). Aplicación de la metodología de aprendizaje cooperativo en materia de Biofarmacia y Farmacocinética. *ARS Pharmaceutica*, 51(2), 261–268.
- Molinero, L. (2008). Meta-análisis, una guía rápida para lectores y usuarios. *Hipertensión*, 25(3), 108–120.
- Morales, P. (1993). Líneas actuales de investigación en métodos cuantitativos: el meta-análisis o la síntesis integradora. *Revista de Educación*, 300, 191–221.
- Moran, D. (2004). The need for evidence-Based Educational Methods. En D. Moran & R. Malott (Ed.) *Evidence-Based Educational Methods* (pp.3–7). California: Elsevier Academic Press.
- Morrison, J., Sullivan, F., Murray, E. & Jolly, B. (1999). Evidence-based education: Development of an instrument to critically appraise reports of educational interventions. *Medical Education* 33, 890–893.
- Moruno, P., Sánchez, M. y Zariquiey, F. (2012). La cultura de la cooperación. El aprendizaje cooperativo como herramienta de diferenciación curricular. En J. Torrego (Coord.). *Alumnos con altas capacidades y aprendizaje cooperativo. Un modelo de respuesta educativa* (pp.167–197). Madrid: Fundación Pryconsa; Fundación SM.
- Muela, C. (2012). Aprendizaje cooperativo de la creatividad publicitaria a través de anuncios audio y/o visuales. *@tic Revista D'Innovació Educativa*, Enero–Junio, 33–39.
- \*Muhammad, Z. (2010). *Effects of cooperative learning intervention on mathematics achievement outcomes and attitudes of non-science college majors* (Tesis doctoral). Southern University and A & M College, USA.
- Mutlu, A. & Eroğ-Tuga, B. (2013). The Role of Computer-Assisted Language Learning (CALL) in Promoting Learner Autonomy. *Eğitim Araştırmaları–Eurasian Journal of Educational Research*, 13(51), 107–122.
- Neber, H., Finsterwald, M. & Urban, N. (2001). Cooperative learning with gifted and high-achieving students: a review and meta-analyses of 12 studies. *High Ability Studies*, 12(2), 199–214.
- Negro, A., Torrego, J. y Zariquiey, F. (2012). Fundamentos del aprendizaje cooperativo. Resultados de las investigaciones sobre su impacto. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.47–73). Madrid: Alianza editorial.



- Nelson, C. (2010). Want Brighter, Harder Working Students? Change Pedagogies! Some Examples, Mainly from Biology. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.119–140). Virginia: Stylus – The National Teaching and Learning Forum.
- Nelson, J. (1986). Improving de lecture through active participation. *College Student Journal*, 20, 316-320.
- Nieto, J. (2004). *Estrategias para mejorar la práctica docente*. Madrid: Editorial CCS.
- \*Nihalani, P., Mayrath, M. & Robinson, D. (2011). When Feedback Harms and Collaboration Helps in Computer Simulation Environments: An Expertise Reversal Effect. *Journal of Educational Psychology*, 103(4), 776–785.
- Ning, H. (2013). The Impact of Cooperative Learning on English as A Foreign Language Tertiary Learners' Social Skills. *Social Behavior and Personality*, 41(4), 557–568.
- Nuhfer, E. (2010). Cooperative Learning in Geological Sciences. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.181–200). Virginia: Stylus – The National Teaching and Learning Forum.
- \*Numpraset, W. (2006). Cooperative Learning and the Achievement of Students in Science, Man, and His Environment. *Au J.T.* 9(3), 139–146.
- Nunnally, J. & Bernstein, I. (1994). *Psychometric Theory* (3ª Ed). New York: McGraw–Hill.
- \*O'Brien, G. & Peters, J. (1994). Effect of four instructional strategies on integrated science process skill achievement of preservice Elementary teachers having different cognitive development levels. *Journal of Elementary Science Education*, 6(1), 30–45.
- O'Donnell, A. (1999). Structuring dyadic interaction through scripted cooperation. En A. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp.179–196). Mahwah, NJ: Lawrence Erlbaum Associates.
- O'Donnell, A. & Dansereau, D. (1995). Scripted Cooperation in Student Dyads: A Methods for Analyzing and Enhancing Academic Learning and Performance. En R. Hertz–Lazarowitz & N. Miller (Ed.). *Interaction in Cooperative Groups. The Theoretical Anatomy of Group Learning* (pp.120–141). United States of America: Cambridge University Press.

- \*O'Donnell, A., Dansereau, D., Hall, R. & Rocklin, T. (1987). Cognitive, social/affective, and metacognitive outcomes of scripted cooperative learning. *Journal of Educational Psychology*, 79(4), 431–437.
- \*O'Donnell, A., Dansereau, D., Hythecker, V., Hall, R., Skaggs, L., Lambiotte, J. & Young, M. (1988). Cooperative procedural learning: effects of prompting and pre- versus distributed planning activities. *Journal of Educational Psychology*, 80(2), 167–171.
- \*O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. & Young, M. (1986). The effects of monitoring on cooperative learning. *Journal of Experimental Education*, 54(3), 169–173.
- \*O'Donnell, A., Dansereau, D., Rocklin, T., Larson, C., Hythecker, V., Young, M. & Lambiotte, J. (1987). Effects of cooperative and individual rewriting on an instruction writing task. *Written Communication*, 4(10), 90–99.
- \*O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J., Larson, C. & Young, M. (1985). Effects of elaboration frequency on cooperative learning. *Journal of Educational Psychology*, 77(5), 572–580.
- O'Donnell, A. & Hmelo–Silver, C. (2013). Introduction: What s Collaborative Learning? An Overview. En C. Hmelo–Silver, C. Chinn, C. Chan & O'Donnell, A. *The International Handbook of Collaborative Learning* (pp.1–15). New York: Routledge.
- \*O'Donnell, A., Larson, C., Dansereau, D. & Rocklin, T. (1986). Effects of cooperation and editing on instruction writing performance. *Journal of Experimental Education*, 54(4), 207–210.
- \*Oh, H. (1988). *The effects of individualistic, cooperative task, and cooperative incentive structures on college achievement in computer programming in BASIC* (Tesis doctoral). Illinois State University, USA.
- Olmos, J. (2012). *Una propuesta para la gestión y evaluación del trabajo cooperativo. Aplicación en el aprendizaje del enlace químico en educación secundaria*. Alemania: Editorial Académica Española.
- Onwuegbuzie, A. & DaRos–Voseles, D. (2001). The role of cooperative learning in research methodology courses: a mixed-methods analysis. *Research in the Schools*, 8(1), 61-75.

- Opdecam, E., Everaert, P., Van Keer, H. & Buysschaert, F. (2014). Preferences for Team Learning and Lecture-Based Learning Among First-Year Undergraduate Accounting Students. *Research Higher Education*, 55, 400–432.
- Organisation for Economic Cooperation and Development (2005). *La definición y selección de competencias clave*. Recuperado de <http://www.deseco.admin.ch/bfs/deseco/en/index/03/02.parsys.78532.downloadList.94248.DownloadFile.tmp/2005.dscexecutivesummary.sp.pdf>
- Ornstein, A. & Scarpaci, R. (2012). *The Practice of Teaching: A Narrative and Case-Study Approach*. United States of America: Waveland Press, Inc.
- Ortega, P., Mínguez, R. y Gil, R. (1997). Aprendizaje cooperativo y desarrollo moral. *Revista Española de Pedagogía*, 206, 33–51.
- Ortego, C. & Botella, J. (2010). The hard but necessary task of gathering order—one effect size indices in meta-analysis. *Psicológica*, 31, 291–315.
- Orwin, R. (1983). A Fail-Safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8(2), 157-159.
- Orwin, R. & Vevea, J. (1994). Identifying potentially interesting variables and analysis opportunities. En H. Cooper, L. Hedges & J. Valentine (Eds.) *The Handbook of Research Synthesis* (2ª Ed.) (pp.177–203). New York: Russell Sage Foundation.
- Ouellett, M. & Fraser, E. (2011). Interracial Team Teaching in Social Work. En K. Plank. *Team Teaching. Across the Disciplines, Across the Academy* (pp.73–96). Virginia: Stylus – The National Teaching and Learning Forum.
- Ovejero, A. (1990a). *El aprendizaje cooperativo. Una alternativa eficaz a la enseñanza tradicional*. Barcelona: Promociones y Publicaciones Universitarias.
- Ovejero, A. (1990b). Las habilidades sociales y su entrenamiento. Un enfoque necesariamente psicosocial. *Psicothema*, 2, 93–112.
- Ovejero, A. (1993). Aprendizaje cooperativo: una eficaz aportación de la psicología social a la escuela del siglo XXI. *Psicothema*, 5, 373–391.
- Ovejero, A., Gutiérrez, M. y Fernández Alonso, J.A. (1996). Eficacia del aprendizaje cooperativo para la integración escolar: Una experiencia en 2º ciclo de E.G.B. *Aula Abierta*, 68, 97–113.
- \*Overlock, T. (1994). *Comparison of Effectiveness of Collaborative Learning Methods and Traditional Methods in Physics Classes at Northern Maine Technical College* (Tesis doctoral). Nova University, USA.
- Pallarés, M. (1990). *Técnicas de grupo para educadores*. Madrid: ICCE.

- Palomares, A. (2011). El modelo docente universitario y el uso de nuevas metodologías en la enseñanza, aprendizaje y evaluación. *Revista Educación*, 355, 591–604.
- Panitz, T. (1999). *Collaborative versus Cooperative Learning: A Comparison of the Two Concepts Which Will Help Us Understand the Underlying Nature of Interactive Learning*. Opinion Papers. Recuperado de <http://eric.ed.gov/?id=ED448443>
- Panitz, T. (2010). Cooperative Learning Structures Help College Students Reduce Math Anxiety and Succeed in Developmental Courses. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.57–68). Virginia: Stylus – The National Teaching and Learning Forum.
- Park, E. & Choi, B. (2014). Transformation of classroom spaces: traditional versus active learning classroom in colleges. *Higher Education*, 68, 749–771.
- Parkay, F., Oaks, M. & Peters, D. (2000). Promoting Group Investigation in a Graduate–Level ITV Classroom. *T.H.E. Journal* 27(9), 86–97.
- Parra–Meroño, M. y Peña–Acuña, B. (2012). El aprendizaje cooperativo mediante actividades participativas. *Anales de Universidad Metropolitana*, 12(2), 15–37.
- Pascarella, E. et al. (2006). Institutional selectivity and good practices in undergraduate education: How strong is the link? *Journal of Higher Education*, 77, 251–285.
- Pascarella, E., Goodman, K., Seifert, T., Tagliapietra–Nicolì, G., Park, S. & Whitt, E. (2007). College student binge drinking and academic achievement: A longitudinal replication and extension. *Journal of College Student Development*, 48(6), 715–727.
- Pascarella, E., Wolniak, G., Seifert, T., Cruce, T. & Blaich, C. (2005). *Liberal arts colleges and liberal arts education: New evidence on impacts*. San Francisco, CA: Jossey–Bass & the Association for the Study of Higher Education.
- Paulson, D. & Faust, J. (2008). *Active Learning for the College Classroom*. Recuperado de <http://www.calstatela.edu/dept/chem/chem2/Active/main.htm>
- \*Pavese, A. (1993). *The effect of cooperative and individual learning on the achievement of at-risk community college students* (Tesis doctoral). Hofstra University, USA.
- Pawson, R. & Tilley, N. (1994). What Works in evaluation research? *British Journal of criminology*, 34, 291–306.

- Pérez, A., Paz, M. y Poveda, P. (2009). Aprendizaje cooperativo y formación de un profesor. Un estudio bibliométrico (1997–2008). *Anales de Documentación*, 12, 209–220.
- Pérez, C. (2003). Cómo desarrollar habilidades sociales mediante el aprendizaje cooperativo. *Aula de Innovación Educativa*, 125, 63–67.
- \*Perihan, A. & Kamuran, T. (2007). The Effectiveness of Jigsaw II on Prospective Elementary School Teachers. *Asia–Pacific Journal of Teacher Education*, 35(2), 129–141.
- Perkins, D. & Saris, R. (2001). A “Jigsaw Classroom” Technique for Undergraduate Statistics Courses. *Teaching of Psychology*, 28(2), 111–113.
- Pértiga, S. y Pita, S. (2005). *Revisiones sistemáticas y meta-análisis (II)*. Recuperado de <https://www.fisterra.com/mbe/investiga/metaanalisis/RSyMetaanalisis2.asp>
- Piaget, J. (1924/1993). Apéndice. Comentarios sobre las observaciones críticas de Vigotsky. En L. Vigotsky. *Pensamiento y Lenguaje* (pp.199–215). Buenos Aires: La Pléyade.
- Piaget, J. (1950). *The psychology of intelligence*. New York: Harcourt.
- Piaget, J. (1983). *El criterio moral en el niño* (4ª Ed.). Barcelona: Fontanella.
- Piaget, J. (1986). *Seis estudios de psicología* (2ª Ed.). Barcelona: Ariel.
- Pigott, T. (2009). Handling missing data.. En H. Cooper, L. Hedges & J. Valentine (Editors). *The handbook of research synthesis and meta-analysis* (2ª Ed.) (pp.399–416). New York: Russell Sage Foundation.
- Pigott, T. (2012). *Advances in meta-analysis*. USA: Springer.
- Plank, K. (Edit.) (2011). *Team Teaching. Across the Disciplines, Across the Academy*. Virginia: Stylus – The National Teaching and Learning Forum.
- Poling, D. & Hupp, J. (2009). Active Learning Through Role Playing. Virtual Babies in a Child Development Course. *College Teaching*, 57(4), 221–226.
- Pons, R., González–Herrero, M. y Serrano, J. (2008). Aprendizaje cooperativo en matemáticas: Un estudio intracontenido. *Anales de Psicología*, 21(2), 253–261.
- Pooper, K. (1991). *Conjeturas y refutaciones. El desarrollo del conocimiento científico* (3ª Ed.). (Trad. N. Miguez). Barcelona: Paidós Básica (Original en inglés, 1972).
- Prais, S.J. (1999). Within–class grouping: A rejoinder. *National Institute Economic Review*, 169, 109–110.
- Program for Large Network Analysis – PAJEK (version 3.15) [Software de computación]. Eslovenia: Pajek.

- Promising practices network (2012). *How programs are considered*. Recuperado de <http://www.promisingpractices.net/criteria.asp>
- Puchau, M. V. (1999). El aprendizaje cooperativo en el aula de música como medida de atención a la diversidad. *Eufonía, Didáctica de la Música*, 17, 115–117.
- Pujolàs, P. (2004). *Aprender juntos alumnos diferentes. Los equipos de aprendizaje cooperativo en el aula*. Barcelona: Eumo–Octaedro.
- Pujolàs, P. (2008). *9 ideas clave: El aprendizaje cooperativo*. Barcelona: Graó.
- Pujolàs, P. (2009). La calidad en los equipos de aprendizaje cooperativo. Algunas consideraciones para el cálculo del grado de cooperatividad. *Revista de Educación*, 349, 225–239.
- Pujolàs, P. (2012). La implementación del aprendizaje cooperativo en las aulas. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.77–104). Madrid: Alianza editorial.
- Qin, Z., Johnson, D. & Johnson, R. (1995). Cooperative Versus Competitive Efforts and Problem Solving. *Review of Educational Research*, 65(2), 129–143.
- Quinn, L. (2006). Investigating Culture Through Cooperative Learning. *Foreign Language Annals*, 39(1), 11–21.
- Rabin, L. & Nutter–Upham, K. (2010). Introduction of a Journal Excerpt Activity Improves Undergraduate Students’ Performance in Statistics. *College Teaching*, 58, 156–160.
- Ramos–Álvarez, M. y Catena, A. (2004). Normas para la elaboración y revisión de artículos originales experimentales en ciencias del comportamiento. *International Journal of Clinical and Health Psychology*, 4(1), 173–189.
- Ramos–Álvarez, M., Moreno–Fernández, M., Valdés–Conroy, B. & Catena, A. (2008). Criteria of the peer review process for publication of experimental and quasi–experimental research in psychology: a guide for creating research papers. *International Journal of Clinical and Health Psychology*, 8(3), 751–764.
- \*Randolph, W. (1992). *The effects of cooperative learning on academic achievement in introductory college biology* (Tesis doctoral). Washington State University, USA.
- Rayón, L. (2012). La puesta en marcha en el aula: analizando la práctica. En J. Torrego y A. Negro (Coords.) *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implementación* (pp.241–268). Madrid: Alianza editorial.

Real Decreto 1791/2010, de 30 de diciembre, por el que se aprueba el Estatuto del Estudiante Universitario. Recuperado de <http://www.boe.es/boe/dias/2010/12/31/pdfs/BOE-A-2010-20147.pdf>

Real Decreto 55/2005, de 21 de enero, por el que se establece la estructura de las enseñanzas universitarias y se regulan los estudios universitarios oficiales de Grado. Recuperado de <http://www.boe.es/boe/dias/2005/01/25/pdfs/A02842-02846.pdf>

\*Reglin, G. L. (1990). The effects of individualized and cooperative computer assisted instruction on mathematics achievement and mathematics anxiety for prospective teachers. *Journal of Research on Computing in Education*, 22, 404–12.

Reimer, T., Reimer, A. & Czienskowski, U. (2010). Decision-Making Groups Attenuate the Discussion Bias in Favor of Shared Information: A Meta-Analysis. *Communication Monographs*, 77(1), 121–142.

Resta, P. & Laferrière, T. (2007). Technology in of Collaborative Learning. *Educational Psychology Review*, 19, 65–83.

Review Manager (RevMan) (Version 5.2) [Computer program]. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012. Recuperado de <http://tech.cochrane.org/revman/about-revman-5>

\*Rewey, K., Dansereau, D., Skaggs, L., Hall, R. & Pitre, U. (1989). Effects of scripted cooperation and knowledge maps on the processing of technical material. *Journal of Educational Psychology*, 81(4), 604–609.

Reyes, E. y Gálvez, J. (2010). Experiencias docentes en innovación educativa como mejora de una enseñanza tradicional en los materiales de construcción. *Formación Universitaria*, 3(4), 13–24.

Richter, A., Dawson, J. & West, M. (2011). The effectiveness of teams in organizations: a meta-analysis. *International Journal of Human Resource Management*, 22(13), 2749–2769.

Richter, R. & Thomas, M. (2011). Arts and Community. En K. Plank. *Team Teaching. Across the Disciplines, Across the Academy* (pp.55–72). Virginia: Stylus – The National Teaching and Learning Forum.

Rico, R., Sánchez-Manzanares, M., Gil, F. & Gibson, C. (2008). Team implicit coordination processes: A team knowledge based approach. *Academy of Management Review*, 23(1), 163–184.

- \*Riggio, R. E., Fantuzzo, J. W., Connelly, S., & Dimeff, L. A. (1991). Reciprocal peer tutoring: A classroom strategy of promoting academic and social integration in undergraduate students. *Journal of Social Behavior and Personality*, 6, 387–396.
- Rittschof, K. & Griffin, B. (2001). Reciprocal Peer Tutoring: re-examining the value of a co-operative learning technique to college students and instructors. *Educational Psychology*, 21(3), 313–331.
- Rivero, L., Samino, R. y Pérez del Campo, E. (2008). *Rendimiento académico y modelos virtuales de enseñanza universitaria en Economía de la Empresa. Nuevos retos hacia la globalización docente en el marco del Espacio Europeo de Educación Superior*. (comunicación). Coord. J. Pindado y G. Payne Estableciendo puentes en una economía global, Vol. 1, 35. XXII Congreso de la Asociación Europea de Dirección y Economía de Empresa, Salamanca.
- Robinson, P. & Cooper, J. (2010). The Interactive Lecture in a Research Methods and Statistics Class. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.105–118). Virginia: Stylus – The National Teaching and Learning Forum.
- Rodríguez, A., Miaja, P., Arias, M., Rodríguez, M. & Lamar, D. (Abril, 2010). *An Experience in Cooperative Learning Developing a Real Aerospacial Project*. IEEE Educon Education Engineering 2010 – The Future of Global Learning Engineering Education (pp.1225–1230). Madrid-España.
- Romero, C. (2009). Cooperative learning instruction & science achievement for secondary and early post-secondary students: a systematic review (Tesis doctoral). Colorado State University.
- \*Roon, R., Van Pilsum, J., Harris, I., Rosenberg, P., Johnson, R., Liaw, C. & Rosenthal, L. (1983). The experimental use of cooperative learning groups in a biochemistry laboratory/course for first-year medical students. *Biochemical Education*, 11(1), 12–15.
- Rosales, C. (2009). Una universidad para el S.XXI. Espacio Europeo de Enseñanza Superior (EEES). Una mirada transdisciplinar, ecoformadora e intercultural. En A. Medina, M.L. Sevillano y S. de la Torre (Coord.) *Innovación y constantes en los estudios de la universidad* (pp.43–51). Madrid: Editorial Universitat, S.A.
- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, 86, 638-641.



- Roseth, C., Johnson, D. & Johnson, R. (2008). Promoting Early Adolescents' Achievement and Peer Relationships: The Effects of Cooperative, Competitive, and Individualistic Goal Structures. *Psychological Bulletin*, 134(2), 223–246
- Roseth, C., Saltarelli, A. & Glass, C. (2011). Effects of Face-to-Face and Computer-Mediated Constructive Controversy on Social Interdependence, Motivation, and Achievement. *Journal of Educational Psychology*, 103(4), 804–820.
- Rysavy, S. & Sales, G. (1991). Cooperative Learning in Computer Based Instruction. *Educational Technology Research and Development*, 139(2), 70–79.
- \*Şahin, A. (2010). Effects of jigsaw II technique on academic achievement and attitudes to written expression course. *Educational Research and Reviews*, 5(12), 777–787.
- Saini, M. & Shlonsky, A. (2012). *Systematic synthesis of qualitative research*. United States of America: Oxford University Press, Inc.
- Sáiz, M. y Román, J. (2011). Cuatro formas de evaluación en Educación Superior gestionadas desde la tutoría. *Revista de Psicodidáctica*, 16(1), 145–161.
- Salas, E., Nichols, D. & Driskell, J. (2007). Testing three team training strategies in intact teams. *Small Groups Research*, 38(4), 471–488.
- Salmerón, H., Rodríguez, S. y Gutiérrez, C. (2010). Metodologías que optimizan la comunicación en entornos de aprendizaje virtual. *Comunicar*, 34, 163–171.
- Sánchez-Meca, J. (2003). La revisión del estado de la cuestión: el meta-análisis. En C. Camisón, M.J. Oltra y M.L. Flor (Eds.), *Enfoques, problemas y métodos de investigación en Economía y Dirección de Empresas. Tomo I* (pp.101-110). Castellón: Universitat Jaume I
- Sánchez-Meca, J. (2010). Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta*, 38(2), 53–64.
- Sánchez-Meca, J. y Ato, M. (1989). Meta-análisis: una alternativa metodológica a las revisiones tradicionales de la investigación. En J. Arnau y H. Carpintero (Eds.) *Tratado de Psicología General I: historia, teoría y método* (pp.617–669). Madrid: Alhambra.
- Sánchez-Meca, J. y Botella, J. (2010). Revisiones sistemáticas y meta-análisis: herramientas para la práctica profesional. *Papeles del Psicólogo*, 31(1), 7–17.
- Sánchez-Meca J., López-Pina J. y Nuñez-Nuñez, R. (2012). Validez y fiabilidad de una escala para la evaluación de la calidad de estudios primarios 95-101.

- En: *Actas Del XII Congreso de Metodología de las Ciencias Sociales y de la Salud*. 978-84-9860-633-1
- Sánchez–Meca, J., Marín, F. y Huedo, T. (2006). Modelo de efectos fijos y modelo de efectos aleatorios. En J. L. Martín, A. Tobías y T. Seoane (Coord.) *Revisiones sistemáticas en las ciencias de la vida. El concepto salud a través de la síntesis de la evidencia científica* (pp.189–204). Recuperado de <http://www.um.es/metaanalysis/pdf/5003.pdf>
- Sánchez–Meca, J., Marín–Martínez, F. y López–López, J. (2011). Meta–análisis e intervención psicosocial basada en la evidencia. *Psychosocial Intervention*, 20(1), 95–107.
- Sánchez, C. (1996). *El movimiento renovador de la Experiencia Somosaguas*. Madrid: Narcea.
- Sánchez, M. y Pueo, B. (2012). Aprendizaje colaborativo mediante corresponsabilidad funcional: una posible solución a las limitaciones habituales en asignaturas de tecnología audiovisual. *Área Abierta*, 33, 1–18.
- Sancho, J. (2001). Docencia e investigación en la universidad: una profesión, dos mundos. *Educación*, 28, 41–60.
- Santos Rego, M. (1990). Estructuras de aprendizaje y métodos cooperativos en educación. *Revista Española de Pedagogía*, 185, 53–78.
- Santos Rego, M. (1999). Aprendizaje cooperativo y rendimiento escolar: balance de perspectivas para la innovación educativa. *Revista Galega do Ensino*, 24, 305–321.
- Santos Rego, M. (2005). La universidad ante el proceso de convergencia europea: un desafío de calidad para la Unión. *Revista Española de Pedagogía*, 230, 5–16.
- Sanz, S. y Reig–Hernández, D. (2013). El aprendizaje social y los profesionales de la información. *El Profesional de la Información*, 22(6), 545–553.
- Savin–Baden, M. (2000). *Problem–based Learning in Higher Education: Untold Stories*. Buckingham: The Society for Research into Higher Education & Open University Press.
- Schoeman, J., Van Schoor, M., Van der Merwe, L. & Meinjes, R. (2009). A case–based, small–group cooperative learning course in preclinical veterinary science aimed at bridging basic science and clinical literacy. *Journal of the South African Veterinary Association*, 80(1), 31–36.

- Schrock, C. (1995). Learning Together in Mathematics. En H. Foyle (Edit.) *Interactive Learning in the Higher Education Classroom. Cooperative, Collaborative, and Active Learning Strategies* (pp.94–103). Washington: The NEA Professional Library Higher Education Series.
- Schroeder, C., Scott, T., Tolson, H., Huang, T & Lee, Y. (2007). A Meta-Analysis of National Research: Effects of Teaching Strategies on Student Achievement in Science in the United States. *Journal of Research in Science Teaching*, 44(10) 1436–1460.
- \*Sears, D. & Pai, H. (2012). Effects of cooperative versus individual study on learning and motivation after reward–removal. *The Journal of Experimental Education*, 80(3), 246–262.
- Sedat, M. (2010). The Effects of Jigsaw IV on the Achievement of Course of Language Teaching Methods and Techniques. *Educational Research and Reviews*, 5(12), 770–776.
- Seifert, T., Goodman, K., King, P. & Baxter, M. (2010). Using Mixed Methods to Study First–Year College Impact on Liberal Arts Learning Outcomes. *Journal of Mixed Methods Research*, 4(3), 248–267.
- Serrano, J. y González–Herrero, M. (1996). *Cooperar para aprender. ¿Cómo implementar el aprendizaje cooperativo en el aula?* Murcia: DM/PPU Eds.
- Serrano, J., González–Herrero, M. y Martínez–Artero, M. (1997). *Aprendizaje cooperativo en matemáticas. Un método de aprendizaje cooperativo–individualizado para la enseñanza de las matemáticas*. Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Serrano, J., González–Herrero, M. y Pons, R. (2008). *Aprendizaje cooperativo en matemáticas. Diseño de actividades en Educación Infantil, Primaria y Secundaria*. Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Serrano, J., Pons, M. y Ruiz, M. (2007). Perspectiva histórica del aprendizaje cooperativo: un largo y tortuoso camino a través de cuatro siglos. *Revista Española de Pedagogía*, 236.
- Serrano, J. y Pons, M. (2014). Influencia del parámetro de mutualidad en la formación del profesorado universitario en y mediante métodos de aprendizaje cooperativo. *Infancia y Aprendizaje*, 36(3), 361–373.
- Serrano, R., Muñoz, J. y López, I. (2012). Valoración por los estudiantes de la licenciatura en psicopedagogía de una experiencia innovadora basada en

- estrategias de aprendizaje colaborativo y el uso de Courselab. *Teoría de la Educación, Educación y Cultura en la Sociedad de la Información*, 13(1), 437–458.
- Shadish, W., Chacón–Moscoso, S. & Sánchez–Meca, J. (2005). Campbell Collaboration and Evidence–Based Decision–Making: enhancing systematic reviews of program evaluation results in Europe. *Evaluation. The International Journal of Theory, Research and Practice*, 11, 95–109.
- Shadish, W., Cook, T. & Campbell, D. (2002). *Experimental and quasi–experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shadish, W. & Myers, D. (2004). *Campbell collaboration research design policy brief*. Recuperado de [http://www.campbellcollaboration.org/artman2/uploads/1/C2\\_Research\\_Design\\_Policy\\_Brief-2.pdf](http://www.campbellcollaboration.org/artman2/uploads/1/C2_Research_Design_Policy_Brief-2.pdf)
- Shadle, S. (2010). Cooperative Learning in General Chemistry Through Process–Oriented Guided Inquiry Learning. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the Disciplines, Across the Academy* (pp.35–56). Virginia: Stylus – The National Teaching and Learning Forum.
- Sharan, S. (1980). Cooperative Learning in Small Groups: Recent Methods and Effects on Achievement, Attitudes, and Ethnic Relations. *Review of Educational Research*, 50(2), 241–271.
- Sharan, Y. & Sharan, S. (1990). Group Investigation Expands Cooperative Learning. *Educational Leadership*, 17–21.
- Sharan, Y. & Sharan, S. (1999). Group Investigation in the Cooperative Classroom. En S. Sharan (Edit.). *Handbook of Cooperative Learning Methods* (pp.97–114). Westport – Connecticut: Praeger.
- Sheng, L. & Xu, J. (Julio, 2009). *A research on Web–based Cooperative Learning of Hearing Impaired University learner*. Proceedings of 2009 4<sup>th</sup> International Conference on Computer Science & Education (pp.1576–1579). Naning, China.
- Sherman, L. (Abril, 1991). *Cooperative Learning in Post Secondary Education: Implications from Social Psychology for Active Learning Experiences*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Singaram, V., Van der Vleuten, C., Stevens, F. & Dolmans, D. (2011). For most of us Africans, we don't just speak'': a qualitative investigation into collaborative

- heterogeneous PBL group learning. *Advance in Health Science Education*, 16, 297–310.
- Siota, M. (2009). *Aplicación de la técnica del Puzzle de Aronson en la docencia de Derecho Financiero y Tributario*. L. Piña (Coord.) VI Jornada metodológica de derecho financiero y tributario Jaime García Añoveros. La Calidad Jurídica de la producción normativa en España (pp.237-243). Madrid: Instituto de Estudios Fiscales.
- Skala, C., Slater, T. & Adams, J. (2000). Qualitative Analysis of Collaborative Learning Groups in Large Enrollment Introductory Astronomy. *Publications of the Astronomical Society of Australia*, 17(2), 185–193.
- Skinner, B. (1954). The science of learning and the art of teaching. *Harvard Educational Review*, 24(1), 86–97.
- Slavin, R. (1977). Classroom Reward Structure: An Analytical and Practical Review. *Review of Educational Research*, 47(4), 633–650.
- Slavin, R. (1980). Cooperative Learning. *Review of educational research*, 50, 315–339.
- Slavin, R. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, 94, 429–445.
- Slavin, R. (1984). Meta-analysis in education: How has it been used? *Educational Researcher*, 13, 6–15.
- Slavin, R. (1985). An introduction to cooperative learning research. En R. Slavin, S. Sharon, S. Kagan, R. Lazarowitz, C. Webb & R. Schmuck (Eds.) *Learning to cooperate, cooperating to learn* (pp.5–15). Plenum Press: New York.
- Slavin, R. (1986). Best evidence synthesis: An alternative to meta-analysis and traditional reviews. *Educational Researcher*, 15, 5–11.
- Slavin, R. (1989a). Research on cooperative learning: consensus and controversy. *Educational Leadership*, 47, 52–54.
- Slavin, R. (1989b). Class size and student achievement: small effects of small classes. *Educational Psychologist*, 24(1), 99–110.
- Slavin, R. (1991). Synthesis of Research on Cooperative Learning. *Educational Leadership*, 48(5), 71–82.
- Slavin, R. (1992a). Research on cooperative learning: consensus and controversy. En A. Goodsell, M. Maher & V. Tinto (Eds). *Collaborative learning: a sourcebook for higher education* (pp.97–99). University Park, PA: National Center on Postsecondary Teaching, Learning and Assessment.

- Slavin, R. (1992b). *Research methods in Education*. (second edition). Needham Heights, MA: Allyn and Bacon.
- Slavin, R. (1995). *Cooperative Learning* (2ª Ed.). Massachusetts: Allyn & Bacon.
- Slavin, R. (1999a). *Aprendizaje cooperativo. Teoría, investigación y práctica*. Libro traducido por Miguel Wald. Argentina: AIQUE
- Slavin, R. (1999b). Student Teams–Achievement Divisions. En S. Sharan (Edit.). *Handbook of Cooperative Learning Methods* (pp.3–19). Westport – Connecticut: Praeger.
- Slavin, R. (1999c). Comprehensive Approaches to Cooperative Learning. *Theory into practice*, 38(2), 74–79.
- Slavin, R. (2008). Perspective on Evidence–based Research in Education. What Works? Issues in Synthesizing Educational Program Evaluations. *Educational Research*, 37(1), 5–14.
- Slavin, R. (2012). *Educational Psychology: Theory and Practice* (10ª Ed.). Boston: Pearson.
- Slavin, R., Cheung, A., Groff, C. & Lake, C. (2008). Effective Reading Programs for Middle and High Schools: A Best–Evidence Synthesis. *Reading Research Quarterly*, 43(3), 290–322.
- Slavin, R. & Fashola, O. (1998). *Show me the evidence! Proven and promising programs for America's schools*. United States of America: Corwin Press, Inc.
- Slavin, R., Lake, C., Davis, S. & Madden, N. (2011). Effective Programs for Struggling Readers: A Best-evidence Synthesis. *Educational Research Review*, 6, 1–26.
- Slavin, R., Lake, C. & Groff, C. (2009). Effective programs in middle and high school mathematics: a best–evidence synthesis. *Review of Educational Research*, 79(2), 839–911.
- Smith, B. & MacGregor, J. (1992). What is Collaborative Learning? En A. Goodsell, M. Maher, V. Tinto, B. Smith & J. MacGregor (Eds). *Collaborative Learning: A Sourcebook for Higher Education* (pp.10–30). University Park, PA: National Center on Postsecondary Teaching, Learning, and Assessment.
- Smith, K., Matusovich, H., Meyers, K. & Mann, L. (2010). Preparing the Next Generation of Engineering Educators and Researchers. Cooperative Learning in the Purdue University School of Engineering Education PhD Program. En B. Millis (Edit.) *Cooperative Learning in Higher Education. Across the*

- Disciplines, Across the Academy* (pp.91–104). Virginia: Stylus – The National Teaching and Learning Forum.
- Smith, M. (1980). Publication bias and meta-analysis. *Evaluation and Education*, 4, 22–24.
- Smith, M. (1984). *A comparison of cooperative and individualistic learning in associate degree nursing students* (Tesis doctoral). University of Minnesota.
- \*Smith, M., Hinckley, C. & Volk, G. (1991). Cooperative Learning in the Undergraduate Laboratory. *Journal of Chemical Education*, 68(5), 413–415.
- Solsona, N. (1999). El aprendizaje cooperativo: una estrategia para la comunicación. *Aula de Innovación Educativa*, 80, 65–67.
- Sorensen, A. & Bernard, S. (2012). Accelerating what works: using qualitative research methods in developing a change package for a learning collaborative. *Joint Commission Journal on Quality and Patient Safety*, 38, 89–95.
- \*Springer, L. (1997). *Relating concepts and applications through structures active learning*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, USA.
- Springer, L., Stanne, M. & Donovan, S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering and technology: a meta-analysis. *Review of Educational Research*, 69, 21–51.
- Stahl, G., Koschmann, T. & Suthers, D. (2006). Computer-ed collaborative learning: An historical perspective. In R. K. Sawyer (Ed.) *Cambridge handbook of the learning sciences* (pp.409–426). Cambridge, UK: Cambridge University Press.
- Recuperado de [http://GerryStahl.net/cscl/CSCL\\_English.pdf](http://GerryStahl.net/cscl/CSCL_English.pdf)
- Stanne, M., Johnson, D. W. & Johnson, R. T. (1999). Social interdependence and motor performance: A meta-analysis. *Psychological Bulletin* 125, 133–154.
- STARD (2003). *Checklist for the reporting of studies of diagnostic accuracy*. Recuperado de <http://www.stard-statement.org/>
- Statistical Package for the Social Sciences – SPSS (version 17) [Software de computación]. Estados Unidos: IBM.
- Stoner, D. (2004). *The effects of cooperative learning strategies on mathematics achievement among middle-grades students* (Tesis doctoral). University of Georgia.

- Summers, J., Bergin, D. & Cole, J. (2009). Examining the relationships among collaborative learning, autonomy, and student incivility in undergraduate classrooms. *Learning and Individual Differences*, 19, 293–298.
- Summers, M. & Volet, S. (2010). Group work does not necessarily equal collaborative learning: evidence from observations and self-reports. *Europe Journal of Psychology Education*, 25, 473–492.
- Sun, S. (2008). *A Comprehensive Review of Effect Size Reporting and Interpreting Practices in Academic Journals in Education and Psychology* (Trabajo de fin de máster). University of Cincinnati.
- Sutton, A. (2004). Publication bias. En H. Cooper, L. Hedges & J. Valentine (Eds.) *The Handbook of Research Synthesis* (2ª Ed.) (pp.435–452). New York: Russell Sage Foundation.
- Svinicki, M. & McKeachie, W. (2014). *McKeachie's teaching tips. Strategies, research and theory for college and university teachers* (14ª Ed.). Belmont, CA: Wadsworth Cengage Learning.
- \*Tanel, Z. & Erol, M. (2008). Effects of cooperative learning on instructing magnetism: analysis of an experimental teaching sequence. *Latin-American Journal of Physics Education*, 2(2), 124–136.
- Tashakkori, A. & Teddlie, C. (2011). *Handbook of Mixed Methods in Social & Behavioral Research* (2ª Ed.). United States of America: Sage Publications.
- Tejero-González, C., Castro-Morera, M. y Balsalobre-Fernández, C. (2012). Importancia del tamaño del efecto. Una ejemplificación estadística con medidas de condición física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 12(48), 715-727. Recuperado de <http://cdeporte.rediris.es/revista/revista48/artimportancia318.htm>
- Thanh, P., Gillies, R. & Renshaw, P. (2008). Cooperative learning and academic achievement of Asian students: A true story. *International Education Studies*, 1, 83–88.
- The Campbell Collaboration (2014). *What helps? What harms? Based on what evidence?* Recuperado de [http://www.campbellcollaboration.org/about\\_us/index.php](http://www.campbellcollaboration.org/about_us/index.php)
- The Cochrane Collaboration (2014). *Trusted evidence. Informed decisions. Better health.* Recuperado de <http://www.cochrane.org/>



- The Evidence for Policy and Practice Information and Co-ordinating Centre (2014). *Systematic reviews*. Recuperado de <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=53>
- Tilwaldi, D., Takahashi, T., Takata, A. & Koizumi, H. (2010). A Cooperative Distance Learning Method Based on Narrow-Band Internet and Its Evaluation. *Electronics and Communications in Japan*, 93(4), 253–263.
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2ª Ed.). Chicago: University of Chicago Press.
- Tinto, V. (1997). Classrooms as communities. Exploring the Educational Character of student persistence. *Journal of Higher Education*, 68(6), 599–623.
- Tjosvold, D. & Field, R. (1986). Effects of concurrence, controversy and consensus on group decision making. *Journal of Social Psychology*, 125, 355–363.
- Tomcho, T. & Foels, R. (2012). Meta-analysis of group learning activities: empirically based teaching recommendations. *Teaching of Psychology*, 39, 159–169.
- Topping, K. (2005). Trends in Peer Learning. *Educational Psychology*, 25(6), 631–645.
- Torrego, J. y Negro, A. (2012). *Aprendizaje cooperativo en las aulas. Fundamentos y recursos para su implantación*. Madrid: Alianza Editorial.
- \*Tran, V. & Lewis, R. (2012). Effects of cooperative learning on students at An Giang University in Vietnam. *International Education Studies*, 5(1), 86–99.
- \*Treisman, P. (1985). *A study of the mathematics performance of black students at the University of California, Berkeley* (Tesis doctoral). University of California, USA.
- Trigueros, C., Rivera, E. y De la Torre, E. (2011). El Chat como estrategia para fomentar el aprendizaje cooperativo. Una investigación en el Prácticum de Magisterio. *Revista de Curriculum y Formación del Profesorado*, 15(1), 195–210.
- Trujillo, F. (2006). Marco teórico. En F. Trujillo, F. y M. Ariza (Edit.) *Experiencias educativas en aprendizaje cooperativo* (pp.15–42). España: Grupo Editorial Universitario.
- Tuan, L. (2013). Infusing Cooperative Learning into an EFL Classroom. *English Language Teaching*, 3(2), 64–77.
- TUNING Educational Structures in Europe (2000). *What is Tuning?* Recuperado de <http://www.unideusto.org/tuningeu/>

- \*Turner, J. (1995). *The effects of cooperative peer study and individual study on the reading comprehension skills of "at-risk" college students enrolled in a developmental reading and study skills course* (Tesis doctoral). The Johns Hopkins University, USA.
- Tynjälä, P. (2008). Perspective into learning at the workplace. *Educational Research Review*, 3(2), 130–154.
- Ulibarri, A. y Fernández, I. (2010). Construyendo un aprendizaje cooperativo e integrador. Experiencia de innovación docente en el marco del EEES. *Metas de Enfermería*, 13(1), 50–55.
- UNESCO (2011). *Clasificación Internacional Normalizada de la Educación CINE*. Recuperado de <http://www.uis.unesco.org/Education/Documents/iscned-2011-sp.pdf>
- University of Glasgow (s.f.a). *Critical appraisal checklist for an article on an educational intervention*. Department of general practice. Recuperado de <http://www.gla.ac.uk/researchinstitutes/healthwellbeing/research/generalpractice/ebp/checklists/#d.en.19536>
- University of Glasgow (s.f.b). *Critical appraisal checklist for a systematic review*. Department of general practice. Recuperado de <http://www.gla.ac.uk/researchinstitutes/healthwellbeing/research/generalpractice/ebp/checklists/#d.en.19536>
- U.S. Department Education (2014). *The Center for Comprehensive School Reform and Improvement*. Recuperado de [http://www.centerforcsri.org/index.php?option=com\\_frontpage&Itemid=1](http://www.centerforcsri.org/index.php?option=com_frontpage&Itemid=1)
- U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003). *Identifying and implementing educational practices ed by rigorous evidence: a user friendly guide*. Recuperado de <http://www2.ed.gov/rschstat/research/pubs/rigoroussevid/index.html>
- Uriarte, J. (1996). Los efectos del autoconcepto académico en los procesos de aprendizaje cooperativo. *Revista de Psicodidáctica* 1, 193–202.
- \*Urion, D. & Davidson, N. (1992). Student achievement in small-group instruction versus teacher-centered instruction in mathematics. *Primus* 2, 257–64.

- Valentine, J. (1994). Judging the quality of primary research. En H. Cooper, L. Hedges & J. Valentine (Editors). *The handbook of research synthesis and meta-analysis* (2ª Ed.) (pp.129–146). New York: Russell Sage Foundation.
- Valentine, J. & Cooper, H. (2003). *Effect size substantive interpretation guidelines: Issues in the interpretation of effect sizes*. Washington, DC: What Works Clearinghouse.
- Valentine, J. & Cooper, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: the study design and implementation assessment device (Study DIAD). *Psychological Methods* 13(2), 130-149.
- \*Valentino, V. (1988). *A study of achievement, anxiety, and attitude toward mathematics in college algebra students using small group interaction methods* (Tesis doctoral). West Virginia University, USA.
- Van Blankenstein, F., Dolmas, D., Van der Vleuten, C. & Schmidt, H. (2013). Relevant priori knowledge moderates the effect of elaboration during small group discussion on academic achievement. *Instructional Science*, 41, 729–744.
- Van der Vleuten, C. & Driessen, E. (2014). What would happen to education if we take education evidence seriously? *Perspective medicine education*, 3(3), 222–232.
- \*Van Wyk, M. (2010). Do Student Teams Achievement Divisions Enhance Economic Literacy? An Quasi-experimental Design. *Journal Social Science*, 23(2), 83–89.
- Vaquero, A. y de Lorenzo, E. (2009). La satisfacción del alumnado en la implementación de metodologías colaborativas de aprendizaje. *Index de Enfermería*, 18(2), 142–146.
- Veerman, A. & Veldhuis-Diermanse, E. (2006). Collaborative Learning through Electronic Knowledge Construction in Academic Education. En A. O'Donnell, C. Hmelo-Silver & G. Erkens (Edit.) *Collaborative Learning, Reasoning, and Technology* (pp.323-355). New York: Routledge.
- Vergara, D. (2012). Una experiencia educativa de aprendizaje cooperativo en la universidad. *Profesorado. Revista de Curriculum y Formación del Profesorado*, 16(2), 387–402.
- Vigotsky, L. (1934/1979). *El desarrollo de los procesos psicológicos superiores*. Barcelona: Crítica.

- Vigostky, L. (1934/1993). *Pensamiento y lenguaje*. Buenos Aires: La Pléyade.
- Villalvilla, M., López, M. y Llamas, F. (1988). Aprendizaje cooperativo: una vía para solucionar el problema de la integración. *Aula abierta*, 52, 115–134.
- Wang, D., Waldman, D. & Zhang, Z. (2014). A Meta–Analysis of Shared Leadership and Team Effectiveness. *Journal of Applied Psychology*, 99(2), 181–198.
- Wang, K. (2012). The impact of nursing students’ chemistry learning performance assessment in Taiwan: competitive versus non–competitive student team achievement division approaches. *Research in Science & Technological Education*, 30(2), 131–149.
- Wang, X. & Hu, Z. (Marzo, 2009). *A Study on Cooperative English Learning in Network Environment*. First International Workshop on Education Technology and Computer Science (pp.578–582). Wuhan, China.
- Warschauer, M. (1996). Computer–assisted language learning: An introduction. En S. Fotos (Ed.) *Multimedia language teaching* (pp.3–20). Tokyo: Logos International. Recuperado de <http://www.gse.uci.edu/faculty/markw/call.html>
- Warschauer, M. & Healey, D. (1998). Computers and language learning: an overview. *Language Teaching*, 31, 57–71.
- Watson, H. & Rees, C. (2008). Meta–analysis of randomized, controlled treatment trials for pediatric obsessive–compulsive disorder. *Journal of Child Psychology and Psychiatry*, 49(5), 489–498.
- Wicherts, J., Bakker, M. & Molenaar, D. (2011). Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results. *Plos one*, 6(11), 1–7.
- Wicherts, J., Borsboom, D., Kats, J. & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61, 726–728.
- Williams, S. & Kelly, G. (2006). Virtual Reflection: What Teachers Say (and Don’t Say) Online About Their Practice. En A. O’Donnell, C. Hmelo–Silver & G. Erkens (Edit.) *Collaborative Learning, Reasoning, and Technology*. (pp.99–125). New York: Routledge.
- Willis, S. et al. (2002). Small–group work and assessment in a PBL curriculum: a qualitative and quantitative evaluation of student perceptions of the process of working in small groups and its assessment. *Medical Teacher*, 24(5), 495–501.
- Wortman, P. & Bryant, F. (1985). School desegregation a black achievement: an integrative review. *Sociological Methods and Research*, 13(3), 289–324.

- Xunta de Galicia (sin fecha). *Meta-análisis*. Recuperado de <http://dxsp.sergas.es/ApliEdatos/Epidat/Ayuda/11-Ayuda%20Meta-an%Elisis.pdf>
- \*Yaeger, P., Marra, R., Costanzo, F. & Gray, G. (Noviembre, 1999). *Interactive Dynamics: Effects of Student-Centered Activities on Learning*. 29th ASEE/IEEE Frontiers in Education Conference, San Juan, Puerto Rico.
- \*Yamarik, S. (2007). Does cooperative learning improve student learning outcomes? *The Journal of Economic Education*, 38(3), 259–277.
- \*Yang, A. (2005). Comparison of the effectiveness of cooperative learning and traditional teaching methods on Taiwanese college students' English oral performance and motivation toward learning (Tesis doctoral). La Sierra University, USA.
- \*Yetter, G., Gutkin, T., Saunders, A., Galloway, A., Sobansky, R. & Song, S. (2006). Unstructured Collaboration Versus Individual Practice for Complex Problem Solving: A Cautionary Tale. *The Journal of Experimental Education*, 74(2), 137–159.
- Young, E. (2011). Onsite Peer Tutoring in mathematics Content Courses for Pre-Services Teachers. En *IUMPST: The Journal* 2. Recuperado de [www.k-12prep.math.ttu.edu](http://www.k-12prep.math.ttu.edu).
- Zabalza, M. (2003). *Competencias docentes del profesorado universitario. Calidad y desarrollo profesional*. Madrid: Narcea.
- Zabalza, M. (2004). *La enseñanza universitaria: el escenario y sus protagonistas*. Madrid: Narcea S.A.
- Zabalza, M. (2007). *Competencias docentes del profesorado universitario: calidad y desarrollo profesional*. Madrid: Narcea S.A.
- Zarei, A. (2012). The Effects of STAD and CIRC on L2 Reading Comprehension and Vocabulary Learning. *Frontiers of Language and Teaching*, 3, 161–173.
- Zhang, J. (2013). Collaboration, Technology, and culture. En C. Hmelo-Silver, C. Chinn, C. Chan & A. O'Donnell (Eds.) *The International Handbook of Collaborative Learning* (pp.495–508). New York: Routledge.

## **Anexos**

**Anexo 1. Glosario de términos métodos cooperativos (inglés-español)**

Términos del inglés	Términos en español
Co-op Co-op	Co-op Co-op
Collaborative Learning and Computers	Aprendizaje colaborativo y ordenadores
Complex Instruction (CI)	Enseñanza compleja
Computer Supported Collaborative Learning (CSCL)	Aprendizaje colaborativo mediado por ordenadores
Computer Supported Collaborative Work (CSCW)	Aprendizaje colaborativo asistido por ordenador
Computer-Assisted Language Learning (CALL)	Aprendizaje de lenguas asistida por ordenador
Computer-Assisted Instruction (CAI)	Instrucción asistida por ordenadores
Computer-Assisted Learning	Aprendizaje asistido por ordenadores
Cooperative Integrated Reading and Composition (CIRC)	Lectura y escritura integrada cooperativa (LEIC)
Cooperative Investigation	Métodos en tareas específicas
Cooperative Review	Revisión cooperativa
Dyadic Methods	Métodos en parejas
Group Investigation (GI)	Investigación grupal (GI)
Informal Methods	Métodos informales
Jigsaw I	Rompecabezas I
Jigsaw II	Rompecabezas II
Jigsaw III	Rompecabezas III
Jigsaw IV	Rompecabezas IV
Jigsaw Technique	Técnica del Jigsaw
Learning Together (LT)	Aprender juntos
Methods Jigsaw	Métodos Jigsaw
Numbered Heads Together	Cabezas numeradas juntas
Peer Collaboration	Tutoría colaborativa
Peer Teaching	Enseñanza entre iguales
Peer Tutoring	Tutoría entre iguales
Reciprocal Peer Tutoring (RPT)	Tutoría recíproca entre iguales
Scripted Cooperation (SC)	Cooperación guiada
Spontaneous Group Discussion	Discusión grupal espontánea
Student Team Learning (STL)	Equipos de aprendizaje
Student Teams-Achievement Divisions (STAD)	Trabajo en equipo-logro individual (TELI)
Team Assited Individualization (TAI)	Individualización ayudada por equipos (IAE)
Team Building	Construcción de equipos
Team Product	Equipo de producción
Teams-Games-Tournament (TGT)	Torneo de juegos por equipos (TJE)
Think-Pair-Share	Pensar-Formar parejas-Compartir

**Anexo 2. Criterios de calidad en organizaciones americanas e inglesas**

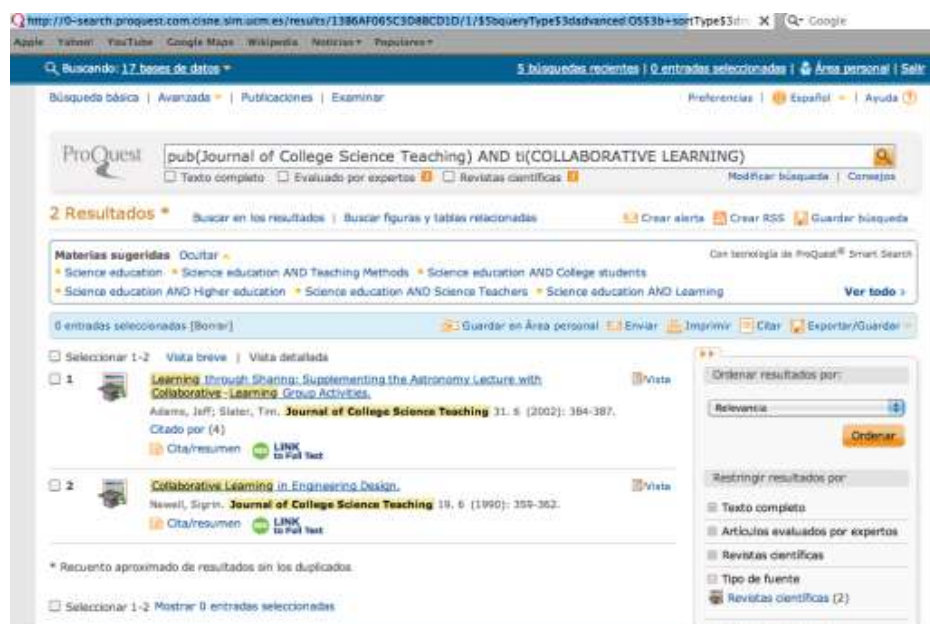
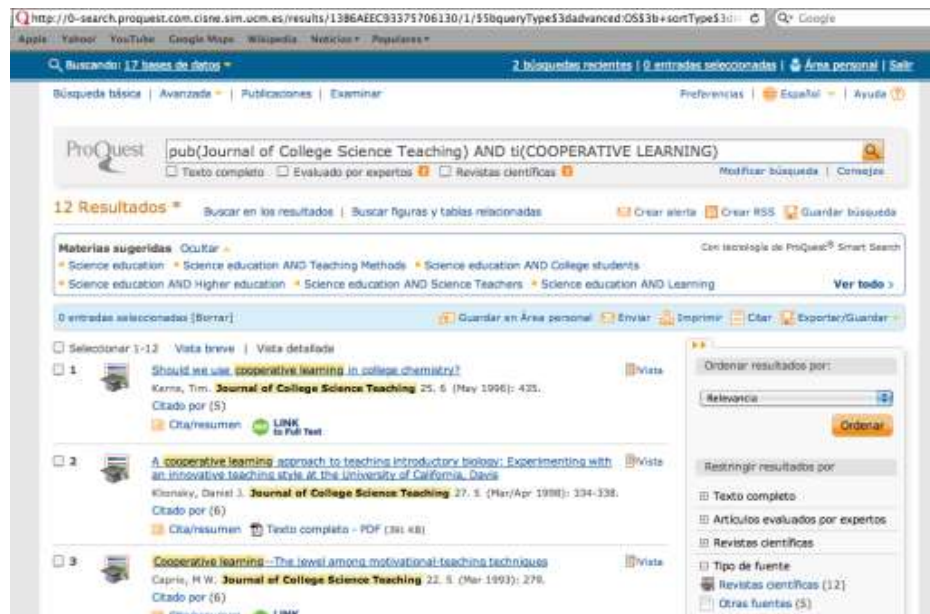
Organización	Lugar de origen y año	Área	¿Qué programas valoran buscando en ellos las mejores evidencias científicas para su posterior síntesis?	¿Cuáles son los criterios de valoración que prevalecen?
What Works Clearinghouse (WWC) <a href="http://ies.ed.gov/ncee/wwc/">http://ies.ed.gov/ncee/wwc/</a>	2002– actual	Educación	Inicio en el proceso de lectura, programas de matemáticas en primaria y ESO, Educación Infantil, programas para el aprendizaje del inglés, prevención de deserción escolar y educación del carácter	Aleatorización diseños experimentales Estudios cuasi-experimentales de calidad son considerados en una categoría de baja rigurosidad
Best Evidence Encyclopedia (BEE) <a href="http://www.bestevidence.org">www.bestevidence.org</a>	2004 –actual	Educación	Programas de matemáticas en primaria, ESO y bachillerato, programas de lectura en ESO y bachillerato, programas de lectura para aprendices del idioma inglés	– Minimizar sesgos – Asignación aleatoria o emparejamiento – Diseños experimentales y cuasi-experimentales – Muestra amplia – Pre-test y pos-test iguales – Duración de la intervención mayor o igual a 12 semanas
Comprehensive School Reform Quality Center (CSRQ) <a href="http://www.csrq.org">www.csrq.org</a>	2003– 2006	Educación	Revisiones sobre los resultados obtenidos de programas integrales orientados a reformas educativas en primaria y secundaria (por ejemplo, Éxito para Todos –aprendizaje cooperativo en la escuela–) y los programas de los proveedores de servicios de educación (por ejemplo escuelas Edison).	Hacen hincapié en el número de estudios y la significación estadística en lugar de evaluaciones aleatorias y los tamaños del efecto
Campbell Collaboration The international <a href="http://www.campbellcollaboration.org">www.campbellcollaboration.org</a>	1994– actual	Ciencias Sociales	Difunde revisiones sistemáticas de la evidencia existente de la investigación en Ciencias Sociales de la educación, la delincuencia, la justicia y el bienestar social. Trabaja para mejorar la metodología de síntesis de la investigación y difundir el estado de la técnica de los métodos de revisión como programas de tutoría después de la escuela	Fuerte énfasis en los ensayos controlados aleatorios
The Cochrane Collaboration <a href="http://www.cochrane.org/">http://www.cochrane.org/</a>	1993 Cuenta con más de 28.000 personas provenientes de 100 países del mundo (universidades, hospitales, fundaciones)	Área de la Salud	Su trabajo se fundamenta en el área de la Salud: prevención, tratamiento y rehabilitación	– Aspectos claves para una revisión sistemática (orientaciones) – Indicaciones claves para la presentación de estudios – Manual Cochrane de revisión sistemáticas de intervenciones en español
Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) <a href="http://www.eppi.ioe.ac.uk">www.eppi.ioe.ac.uk</a>	Universidad de Londres 1993–actual	Salud y Educación	Programas en Ciencias de la Educación, la enseñanza del Inglés y Educación para la ciudadanía. Se centran en variables como por ejemplo, los efectos de la enseñanza de la gramática en la escritura y no en programas específicos	Los criterios de evaluación varían en cuanto a enfoque y metodología dependiendo del área educativa que revisen

Fuente: elaboración propia a partir de Slavin (2008) y la revisión de páginas web de las organizaciones mencionadas



### Anexo 3. Evidencias búsqueda de fuentes primarias

Journal of College Science Teaching



## Higher Education Research

Search results for: **pub(Higher Education Research Development) AND all(COLLABORATIVE LEARNING)**

10 Resultados \*

**Materias sugeridas:** Ocultar

- Higher education AND Learning
- Higher education AND Service Learning
- Higher education AND Learning Strategies
- Higher education AND Learning Disabilities
- Higher education AND Adult Learning
- Higher education AND Learning Activities

**Entradas seleccionadas (Borre):**

- Impact of a New Integrated Medicine Program on Students' Approaches to Learning**

Sakamoto, Chihaya; Davith, Hughes, Christopher; Taskaya, Izzan. **Higher Education Research and Development** 28, 3 (June 2010): 289-302.

...design on student approaches to **learning**. Although the new program was ...promote deeper approaches to **learning**, the results revealed a more ...for both independent and **collaborative learning**—were likely to persist.

Citado por (2) Referencias (31)

Ver resumen [LINK to Full Text](#)
- Measuring Effectiveness of Health Education in a Web-Based Learning Environment: A Preliminary Report**

Luckey, Lari; Patterson, John; Harper, Barry. **Higher Education Research and Development** 18, 3 (1999): 233-246.

Ordenar resultados por: Relevancia

Restringir resultados por: Artículos evaluados por expertos, Revistas científicas, Tipo de fuente

Search results for: **pub(Higher Education Research Development) AND all(COOPERATIVE LEARNING)**

16 Resultados \*

**Materias sugeridas:** Ocultar

- Higher education AND Cooperative Learning
- Higher education AND Cooperative education
- Higher education AND Learning
- Higher education AND Cooperative Planning
- Higher education AND Cooperative Programs
- Higher education AND Service Learning

**Entradas seleccionadas (Borre):**

- Relating Student Learning and Commercial Outcomes in the Workplace**

Lee, Geoffrey; McEvoy, Robert; Holland, Barbara. **Higher Education Research and Development** 26, 5 (October 2010): 563-574.

...growing across higher education, **Cooperative** education and other internships ...experiential and connected **learning**. **Cooperative** programs combine business, ...in-the-job. Most studies of work-based **learning** focus on individual or small

Citado por (1) Referencias (59)

Ver resumen [LINK to Full Text](#)
- Interactive Teaching Goes to Uni: Keeps Students Awake and Learning Alert**

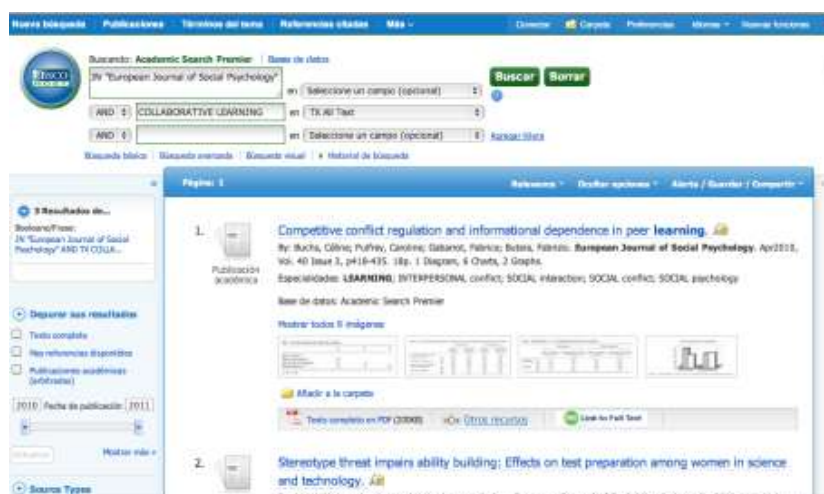
Harrison, Perry. **Higher Education Research and Development** 13, 1 (1994).

...education relies on human ability and **learning** at the Queensland University of Technology.

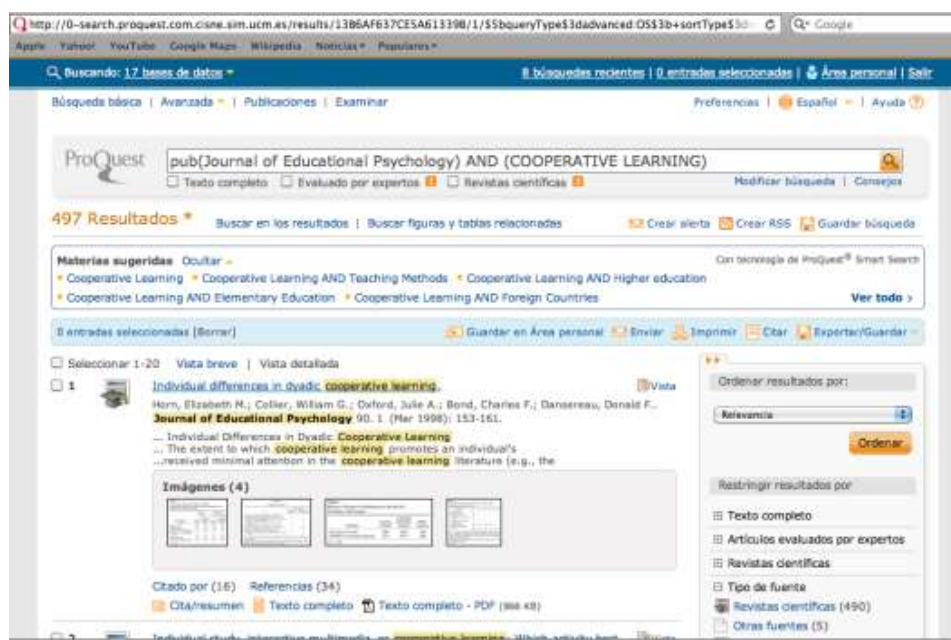
Ordenar resultados por: Relevancia

Restringir resultados por: Artículos evaluados por expertos, Revistas científicas, Tipo de fuente

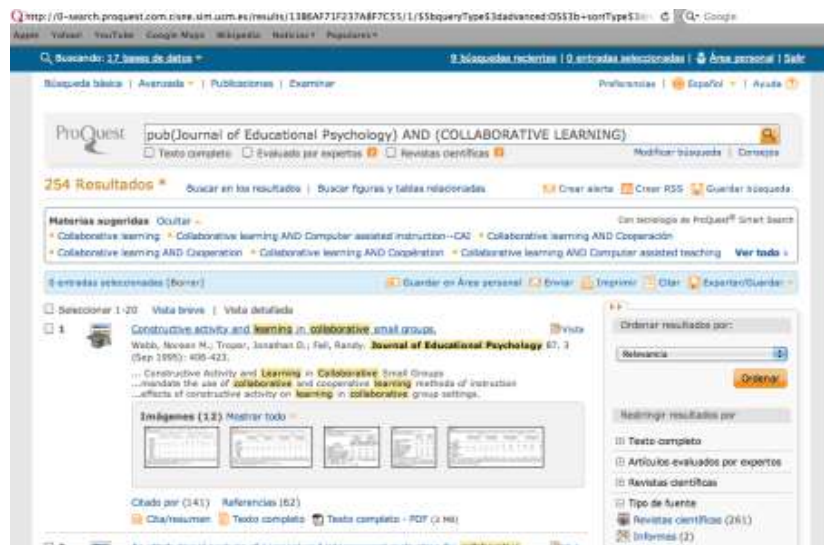




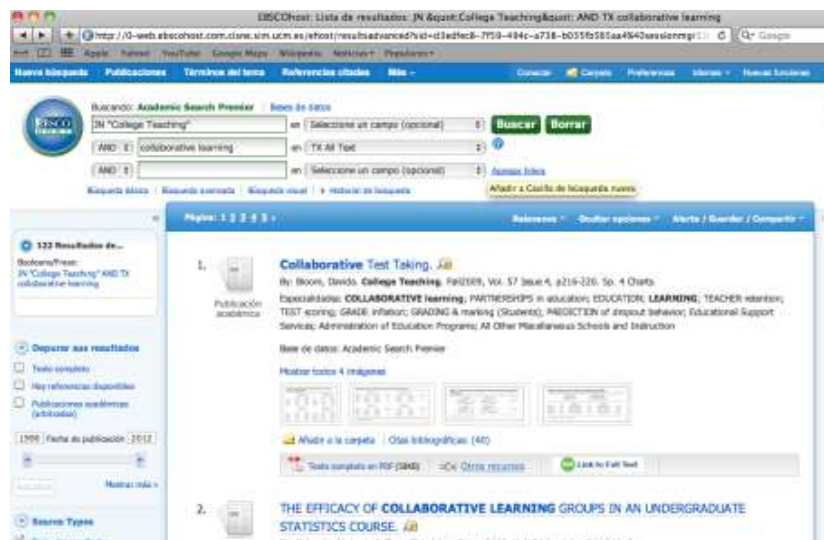
Journal of Educational Psychology

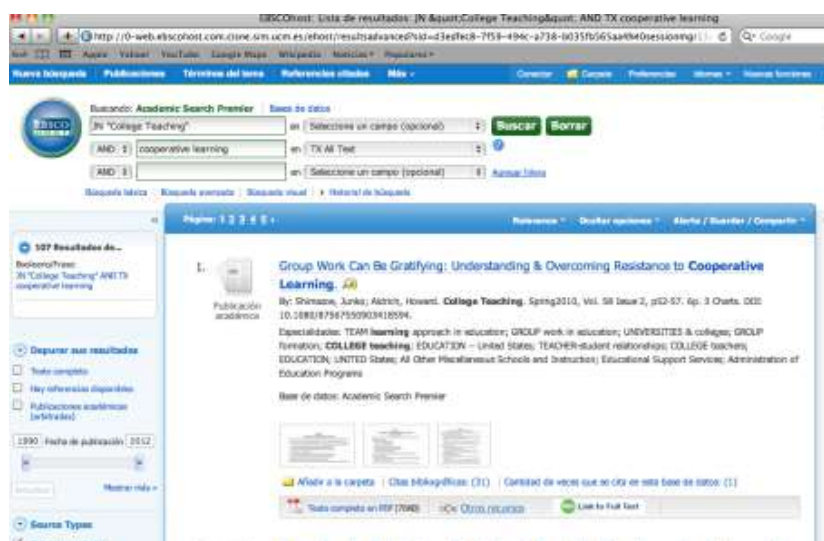




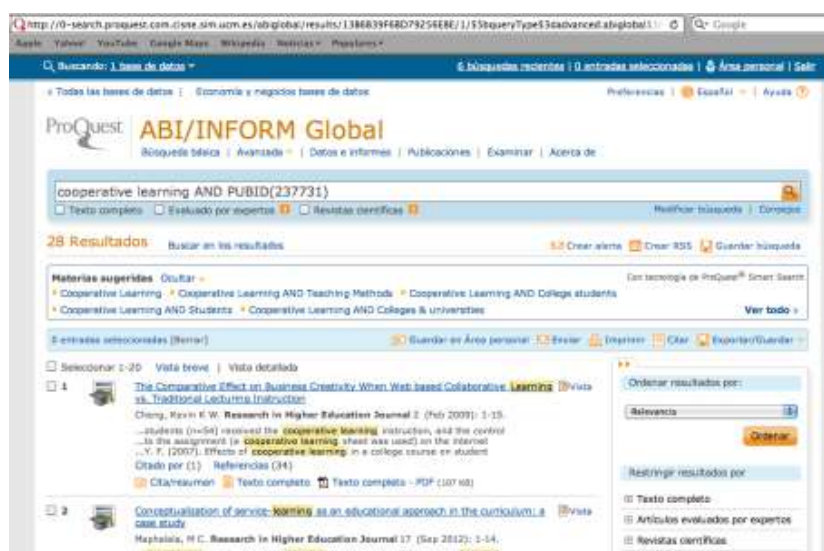


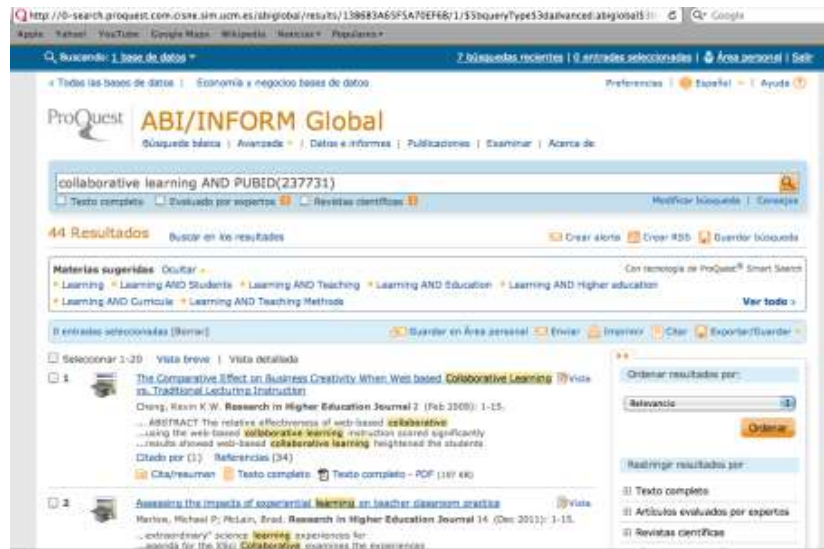
## College Teaching



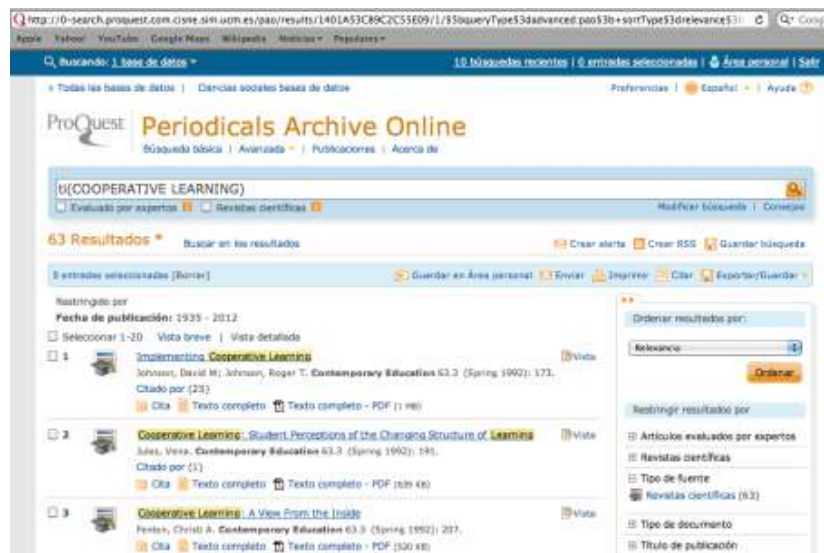


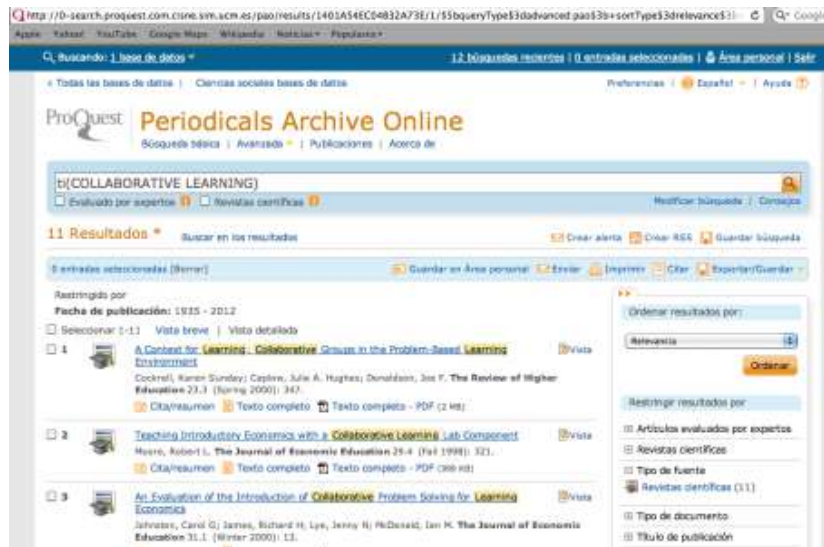
## Research in Higher Education





## Periodicals Archive Online







#### **Anexo 4. Listado estudios del meta-análisis de Bowen 2000**

El listado de referencias facilita la identificación de los estudios primarios utilizados en el meta-análisis de Bowen (2000)

- Banerjee, A. & Vidyapati, T. (1997). Effect of lecture and cooperative learning strategies on achievement in chemistry in undergraduate classes. *International Journal of Science Education*, 19(8), 903–910.
- Basili, P.A. (1988). *Conceptual change strategies within cooperative groups of community college chemistry students: An experiment*. (Tesis Doctoral), University of Maryland.
- Bowen, C. & Phelps, A. (1997). Demonstration-based cooperative testing in general chemistry: A broader assessment-of-learning technique. *Journal of Chemical Education*, 74, 715-719.
- Burron, C.; James, M. & Ambrosio, A. (1993). The effects of cooperative learning in a physical science course for elementary/middle level preservice teachers. *Journal of Research in Science Teaching*, 30, 697–707.
- Dinan, F. & Frydrychowski, V. (1995). A Team Learning Method for Organic Chemistry. *Journal of Chemical Education*, 72 (5), 429-431.
- Dougherty, R. (1997). Grade/Performance Contracts, Enhanced Communication, Cooperative Learning and Student Performance in Undergraduate Organic Chemistry. *Journal of Chemical Education*, 74 (6), 722-726.
- Lunderberg, M. (1990). Supplemental instruction in chemistry. *Journal of Research in Science Teaching*, 27(2), 145–155.
- Metz, P. (1987). *The effects of interactive instruction and lectures on the achievement and attitudes of chemistry students*. (Tesis doctoral), Purdue University.
- Niaz, M. (1995) Cognitive conflict as a teaching strategy in solving chemistry problems: A dialectic–constructivist perspective. *Journal of Research in Science Teaching*, 32, 959–970.
- Ross, M. & Fulton, R. (1994). Active Learning Strategies in the Analytical Chemistry Classroom. *Journal of Chemical Education*, 71(2), 141.
- Smith, M, Hinckley, C. & Volk, G. (1991). Cooperative learning in the undergraduate laboratory. *Journal of Chemical Education*, 68(5), 413–415.

### Anexo 5. Listado estudios del meta-análisis de Springer, Stanne y Donovan 1999

El listado de referencias facilita la identificación de los estudios primarios utilizados en el meta-análisis de Springer, Stanne y Donovan (1999)

- Baker, L. J. (1995). The effect of cooperative study groups on achievement of college-level computer science programming students. (Tesis doctoral), University of Texas at Austin.
- Bonsangue, M. (1991). *Achievement effects of collaborative learning in introductory statistics: A time series residual analysis*. Trabajo presentado The Joint Annual Meeting of the Mathematical Association of America/The American Mathematical Society. San Francisco.
- Bonsangue, M. (1994). An efficacy study of the calculus workshop model. *Research in Collegiate Mathematics Education*, 1(1), 1–19.
- Borresen, C. R. (1990). Success in introductory statistics with small groups. *College Teaching*, 38(1), 26–82.
- DeClute, J. & Ladyshevsky, R. (1993). Enhancing clinical competence using a collaborative clinical education model. *Physical Therapy*, 73(10), 683–97.
- Dees, R. L. (1991). The role of cooperative learning in increasing problem-solving ability in a college remedial course. *Journal for Research in Mathematics Education*, 22(5), 409–21.
- Frierson, H. T. (1986). Two intervention methods: Effects on groups of predominantly black nursing students' board scores. *Journal of Research and Development in Education*, 19(3), 18–23.
- Ganter, S. L. (1994). The importance of empirical evaluations of mathematics programs: A case from the calculus reform movement. *Focus on Learning Problems in Mathematics*, 16(2), 1–19.
- Giraud, G. (1996). *Cooperative learning and statistics instruction*. Manuscrito no publicado, University of Nebraska at Lincoln.
- Hall, D. A. (1992). The influence of an innovative activity-centered biology program on attitudes toward science teaching among preservice elementary teachers. *School Science & Mathematics*, 95(5), 239–242.
- Hanshaw, L. G. (1982). Test anxiety, self-concept, and the test performance of students paired for testing and the same students working alone. *Science Education*, 66(1), 15–24.
- Harding, R. F. & Fletcher, R. K. (1994). Effectiveness of variations in collaborative cooperative learning in RDS mathematics classes. Trabajo presentado en The Annual Meeting of the Tennessee Academy of Science, Nashville.
- Iwasiw, C. L. & Goldenberg, D. (1993). Peer teaching among nursing students in the clinical area: Effects on student learning. *Journal of Advanced Nursing*, 18, 659–668.
- Jimison, L. D. (1990). *A study to investigate the effect of cooperative group learning on selected cognitive and affective outcomes for preservice elementary teachers in a mathematics methods class* (Tesis doctoral), Oklahoma State University.
- Springer, L. (abril de 1997). *Relating concepts and applications through structures active learning*. Trabajo presentado en The Annual Meeting of the American Educational Research Association, Chicago, USA.
- Jones, D. J. & Brickner, D. (1996). *Implementation of cooperative learning in a large-enrollment basic mechanics course*. Trabajo presentado en American Society for Engineering Education Annual Conference Proceedings.
- Kacer, B.; Rocklin, T. & Weinholtz, D. (1990). *Individual versus small group instruction of computer applications: A quantitative and qualitative comparison*. Unpublished manuscript.
- Keeler, C. & Anson, R. (1995). An assessment of cooperative learning used for basic computer skills instruction in the college classroom. *Journal of Educational Computing Research*, 12(4), 379–393.
- Keeler, C. & Steinhurst, R. K. (1994). Cooperative learning in statistics. *Teaching Statistics*, 16(3), 81–84.
- Keeler, C. M. & Voxman, M. (1994). The effect of cooperative learning in remedial freshmen level mathematics. *The AMATYC Review*, 16(1), 37–44.

- Koch, L.-C. (1992). Revisiting mathematics. *Journal of Developmental Education*, 16(1), 12–18.
- Lovelace, T. & McKnight, C. (1980). The effects of reading instruction on calculus students' problem solving. *Journal of Reading*, 23(4), 305–308.
- Lynch, B. (1984). Cooperative learning in interdisciplinary education for the allied health professions. *Journal of Allied Health*, 13(2) 83–93.
- Mehta, J. (1993). Cooperative learning in computer programming at the college level, (Tesis Doctoral), University Of Illinois at Chicago.
- O'Brien, G. & Peters, J. (1994). Effect of four instructional strategies on integrated science process skill achievement of preservice elementary teachers having different cognitive levels. *Journal of Elementary Science Education*, 6(1), 30–45.
- Overlock, T. H. (1994). *Comparison of effectiveness of collaborative learning methods and traditional methods in physics classes at Northern Maine Technical College*. East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED 367 394).
- Pisani, A. M. (1994). Involvement through cooperative learning: An attempt to increase persistence in the biological sciences, (Tesis doctoral), University Of Illinois at Chicago.
- Randolph, W. M. (1992). The effect of cooperative learning on academic achievement in introductory college biology, (Tesis Doctoral), Washington State University.
- Reglin, G. L. (1990). The effects of individualized and cooperative computer assisted instruction on mathematics achievement and mathematics anxiety for prospective teachers. *Journal of Research on Computing in Education*, 22(4), 404–412.
- Shearn, E. & Davidson, N. (1989). *Use of small-group teaching and cognitive developmental instruction in a mathematics course for prospective elementary school teachers*. Trabajo presentado The Annual Meeting of the American Educational Research Association, San Francisco.
- Smith, M. (1984). *A comparison of cooperative and individualistic learning in associate degree nursing students*, (Tesis doctoral), University of Minnesota.
- Treisman, P. (1985). *A study of the mathematics performance of black students at the University of California, Berkeley*, (Tesis doctoral), University of California Berkeley.
- Urión, D. & Davidson, N. (1992). Student achievement in small-group instruction versus teacher-centered instruction in mathematics. *Primus*, 2(3), 257–64.
- Valentino, V. (1988). *A study of achievement, anxiety, and attitude toward mathematics in college algebra students using small-group interaction methods* (Tesis doctoral), West Virginia University.
- Basili, P. (1988). *Conceptual change strategies within cooperative groups of community college chemistry students: An experiment*, (Tesis doctoral), University of Maryland).
- Frierson, H. (1987). Academic performance in predominantly black nursing classes: Effects associated with intervention designed for standardized test preparation. *Journal of Research and Development in Education*, 20, 37–40.
- Johnson, S. & Fischbach, R. (1992). *Teaching problem solving and technical mathematics through cognitive apprenticeship at the community college level*. (Report –NCRVE–MDS– No. 468). Berkeley, CA: National Center for Research on Vocational Education. (ERIC Document Reproduction Services No. ED 352455)
- Lundeberg, M. (1990). Supplemental instruction in chemistry. *Journal of Research in Science Teaching*, 27(2), 145–155.
- Smith, M., Hinckley, C. & Volk, G. (1991). Cooperative learning in the undergraduate laboratory. *Journal of Chemical Education*, 68(5), 413–415.

## Anexo 6. Listado estudios del meta-análisis de Johnson y Johnson 1989

El listado de referencias que se presenta se encuentra solo en el libro de Johnson y Johnson del año 1989 titulado *Cooperation and competition. Theory and research*.

- Abramson, J. (1977). *The effects of non-competitive, individual competitive, and group competitive situations on the verbal and figural creativity of college students* (Tesis doctoral), Michigan State University.
- Almack, J. (1930). Mental efficiency of consulting pairs. *Educational Research Bulletin*, 9, 2–3.
- Andrews, J. (1981). Teaching format and student style: Their interactive effects on learning. *Research in Higher Education*, 14(2), 161–178.
- Barnlund, D. (1959). A comparative study of individual, majority, and group judgment. *Journal of Abnormal and Social Psychology*, 58, 55–60.
- Barrall, M. & Axelrod, S. (1978). The effects of learning partners and retests on pretest–posttest scores, final course grades, and student attitudes. *Research in Higher Education*, 8, 177–187.
- Beach, L. (1960). Sociability and academic achievement in various types of learning situations. *Journal of Educational Psychology*, 51(4), 208–212.
- Beach, L. (1974). Self-directed student groups and college learning. *Higher Education*, 3, 187–200.
- Beaman, A., Diener, E., Fraser, S. & Endresen, K. (1977a). Effect of voluntary and semivoluntary peer monitoring programs on academic performance. *Journal of Educational Psychology*, 69(2), 109–114.
- Bender, T. (1986). Monitoring and the transfer of individual problem solving. *Contemporary Educational Psychology*, 11, 161–169.
- Bergan, J. (1967). The effects of observation and competition on performance on a programmed text. *Psychology in the Schools*, 4(1), 27–29.
- Bodine, R. (1977, April). The effects of cognitive style, task structure, and task setting on student outcomes – cognitive and affective. Trabajo presentado en The Annual Meeting of the American Educational Research Association, New York. (Eric Document Reproduction Service No. ED 135 781).
- Bruning, J. & Mettee, D. (1966). The effects of various social factors on motivation in a competitive situation. *The Journal of Social Psychology*, 70, 295–297.
- Bruning, J. Sommer, D. & Jones, B. (1966). The motivational effects of cooperation and competition in the means–independent situation. *Journal of Social Psychology*, 68, 269–274.
- Carrier, C. & Sales, G. (1987). Pair versus individual work on the acquisition of concepts in a computer–based instructional lesson. *Journal of Computer Based Instruction*, 14(1), 11–17.
- Church, R. (1962a). The effects of competition on reaction time and palmar skin conductance. *Journal of Abnormal and Social Psychology*, 65(1), 32–40.
- Clement, D. (1971). Learning and retention in student–led learning and discussion groups. *Journal of Social Psychology*, 84, 279–286.
- Clifford, J. (1981). Composing in stages: The effects of collaborative pedagogy. *Research in the Teaching of English*, 15(1), 37–53.
- Cook, M. (1982). *Cooperative and competitive intergroup relations and their effects on the processes and outcomes of intragroup cooperation and competition*, (Tesis doctoral), University of Iowa.
- Cook, S. (1971). *The effect of unintended interracial contact upon racial interaction and attitude change*. (Informe de Investigación No. 5–1320). Washington, D.C.: Office of Education (DHEW).
- Cottrell, N. (1964). Means–interdependence, prior acquaintance, and emotional tension during cooperation and subsequent competition. *Human Relations*, 16, 249–262.
- Covington, M. & Omelich, C. (1984). Task–oriented versus competitive learning structures: Motivational and performance consequences. *Journal of Educational Psychology*, 76(6), 1038–1050.
- Crannell, C., Switzer, S. & Morrisette, J. (1965). Individual performance in cooperative and independent groups. *Journal of General Psychology*, 73, 231–236.
- Cratty, B. & Sage, J. (1964). Effect of primary and secondary group interaction upon improvement in a complex movement task. *Research Quarterly*, 35(3), 265–274.
- Crawford, T., & Sidowski, J. (1964). Monetary incentive and cooperation/competition instructions in a minimal social situation. *Psychological Reports*, 15, 233–234.
- Crombag, H. (1966). Cooperation and competition in means–interdependent trades: A replication. *Journal of Personality and Social Psychology*, 4(6), 692–695.
- Davidson, N. (1971). The small group–discovery method as applied in calculus instruction. *American Mathematical Monthly*, 78(7), 789–791.

- Davis, J. (1969). Individual-group problem solving, subject preference and problem type. *Journal of Personality and Social Psychology*, 13(4), 362-374.
- DeCharms, R. (1957). Affiliation motivation and productivity in small groups. *Journal of Abnormal and Social Psychology*, 55, 222-226.
- Dees, R. (1983, April). *The role of co-operation in increasing mathematics problem-solving ability*. Trabajo presentado en The Annual Meeting of the American Educational Research Association, Montreal.
- Deutsch, F. & Leong, F. (1983). Male responses to female competence. *Sex Roles*, 9(1), 79-91.
- Deutsch, M. (1949). An experimental study of the effects of co-operation and competition upon group process. *Human Relations*, 2, 199-231.
- Dumn, R. E., & Goldman, M. (1966). Competition and noncompetition in relationship to satisfaction and feelings towards own- group and nongroup members. *Journal of Social Psychology*, 68, 229-311.
- Durling, R., & Shick, C. (1976). Concept attainment by pairs and individuals as a function of vocalization. *Journal of Educational Psychology*, 68(1), 83-91.
- Faust, W. (1959a). Group versus individual problem-solving. *Journal of Abnormal and Social Psychology*, 59, 68-72.
- Fraser, S., Beaman, A., Diener E., & Kelem, R. (1977). Two, three, or four heads are better than one: Modification of college performance by peer monitoring. *Journal of Educational Psychology*, 69, 101-109.
- Frierson, H. (1987). Academic performance in predominantly Black nursing classes: Effects associated with intervention designed for standardized test preparation. *Journal of Research and Development in Education*, 20(3), 37-40.
- Frost, D. (1977). *The effects of cooperation and competition on the creative expression of college students*, (Tesis doctoral), University of Georgia.
- Goldman, M. (1965). A comparison of individual and group performance for varying combinations of initial ability. *Journal of Personality and Social Psychology*, 1(3), 210-216.
- Goldman, M. (1966). A comparison of group and individual performance where subjects have varying tendencies to solve problems. *Journal of Personality and Social Psychology*, 3(5), 604-607.
- Goldman, M., Dietz, D. & McGlynn, A. (1968). Comparison of individual and group performance related to heterogeneous-wrong responses, size and patterns of interaction. *Psychological Reports*, 23, 459-465.
- Goldman, M. , McGlynn, A., & Toledo, A. (1967). Comparison of individual and group performance of size three and five with various initially right and wrong tendencies. *Journal of Personality and Social Psychology*, 7(2), 222-226.
- Goldman, M., Stockbauer, J. & McAuliffe, T. (1977). Intergroup and intragroup competition and cooperation. *Journal of Experimental and Social Psychology*, 13, 81-88.
- Goldstein, G. (1981). The effects of competition and external rewards on intrinsic motivation (Tesis doctoral), University of New Hampshire.
- Gordon, K. (1924). Group judgments in the field of lifted weights. *Journal of Experimental Psychology*, 7(6), 398-400.
- Grossack, M. (1953). Some effects of cooperation and competition upon small group behavior. *Journal of Abnormal and Social Psychology*, 49, 332-339.
- Gurnee, H. (1937). Maze learning in the collective situation. *Journal of Psychology*, 3, 437-443.
- Gurnee, H. (1968). Learning under competitive and collaborative sets. *Journal of Experimental Social Psychology*, 26-34.
- Haines, D. & McKeachie, W. (1967). Cooperative versus competitive discussion methods in teaching introductory psychology. *Journal of Educational Psychology*, 58(6), 386-390.
- Haley, W. E., & Strickland, B. R. (1986). Interpersonal betrayal and cooperation: effects on self-evaluation in depression. *Journal of Personality and Social Psychology*, 50(2), 386-391.
- Hall, R., Rocklin, T., Dansereau, D., Skaggs, L., O'Donnell, A., Lambiotte, J. & Young, M. (1988). The role of individual differences in the cooperative learning of technical material. *Journal of Educational Psychology*, 80(2), 172-178.
- Hammond, L., & Goldman, M. (1960). Competition and non-competition and its relationship to individual and group productivity. *Sociometry*, 24(1), 46-60.
- Horner, M. (1974). Performance of men in noncompetitive and interpersonal competitive achievement-oriented situations. En J. Atkinson & J. Raynor (Eds.), *Motivation and Achievement* (pp. 237-254). Washington, DC: Winston.
- Hovey, D., Gruber, H. & Terrell, G. (1963). Effects of self-directed study on course achievement retention, and curiosity. *Journal of Educational Research*, 56(7), 346-351.

- Husband, R. (1940). Cooperation versus solitary problem situation. *Journal of Social Psychology*, 11, 405–409.
- Hwong, N., Caswell, A., Johnson, D., & Johnson, R. (1988). *Effects of cooperative and individualistic learning structures on prospective elementary teachers' music achievements and attitudes*. Manuscrito no publicado, University of Minnesota, Minneapolis.
- Hythecker, V., Dansereau, D., O'Donnell A., Larson, C., Lambiotte, J., Rockling, T., & Young, M., (abril de 1985). *The development and evaluation of a strategy for procedure learning*. Trabajo presentado en The Annual Meeting of the Southwestern Psychological Association, Austin, TX.
- Hythecker, V., Rockling, T., Dansereau, D., Lambiotte, J., Larson, C., & O'Donnell, A. (1985a). A computer based learning strategy training module: development and evaluation. *Journal of Educational Computing Research*, 1(3), 275–283.
- Jackson, J., & Williams, K. (1985). Social loafing on difficult tasks: Working collectively can improve performance. *Journal of Personality and Social Psychology*, 49(4), 937–942.
- Johnson, D. & Johnson, R. (1972). Effects of cooperative, competitive, and individualistic conditions on children's problem-solving performance. *American Educational Research Journal*, 17(1), 83–93.
- Johnson, H. & Torcivia, J. (1967). Group and individual performance on a single-stage task as a function of distribution of individual performance. *Journal of Experimental Social Psychology*, 3, 266–273.
- Johnson, R., Bjorkland, R. & Krotee, M. (1984). The effects of cooperative, competitive and individualistic student interaction patterns on the achievement and attitudes of the golf skills of putting. *Research Quarterly*, 55(2), 129–134.
- Johnson, S., & Johnson, D. (1972). The effects of other's actions, attitude similarity, and race on attraction towards others. *Human Relations*, 25(2), 121–130.
- Jones, S. & Vroom, V. (1964). Division of labor and performance under cooperative and competitive conditions. *Journal of Abnormal and Social Psychology*, 68(3), 313–320.
- Julian, J. & Perry, F. (1967). Cooperation contrasted with intra-group and inter-group competition. *Sociometry*, 30(1), 79–90.
- Kelly, R., Rawson, H. & Terry, R. (1973). Interaction effects on achievement need and situational press on performance. *Journal of Social Psychology*, 8, 439–451.
- Lambiotte, J., Dansereau, D., O'Donnell, A., Young, M., Skaggs, L., & Hall, R. (1988). Effects of cooperative script manipulations on initial learning and transfer. *Cognition and Instruction*, 5(2), 103–121.
- Lambiotte, J., Dansereau, D., Rocklin, T., Fletcher, B., Hythecker, V., Larson, C. & O'Donnell, A. (1987). Cooperative Learning and test taking: Transfer of skills. *Contemporary Educational Psychology*, 12(1), 52–61.
- Larson, C., & Dansereau, D. (1983, April). *Cooperative learning: The role of individual differences*. Trabajo presentado en The Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Larson, C., Dansereau, D., Hythecker, V., O'Donnell, A., Young, M., Lambiotte, J. & Rocklin, T. (1986). Technical training: An application of a strategy for learning structural and functional information. *Contemporary Educational Psychology*, 11(3), 217–228.
- Laughlin, P. (1965). Selection strategies in concept attainment as a function of number of persons and stimulus display. *Journal of Experimental Psychology*, 70(3), 323–327.
- Laughlin, P. (1972). Selection versus reception concept-attainment paradigms for individuals and cooperative pairs. *Journal of Educational Psychology*, 63(2), 116–122.
- Laughlin, P. & Adamopoulos, J. (1980). Social combination processes and individual learning for six-person cooperative groups on an intellectual task. *Journal of Personality and Social Psychology*, 38(6), 941–947.
- Laughlin, P. & Barth, J. (1981). Group-to-individual and individual-to-group problem-solving transfer. *Journal of Personality and Social Psychology*, 41(6), 1087–1093.
- Laughlin, P. & Bitz, D. (1975). Individual versus dyadic performance on a disjunctive task as a function of initial ability level. *Journal of Personality and Social Psychology*, 31(3), 487–496.
- Laughlin, P. & Branch, L. (1972). Individual versus tetradic performance on a complementary task as a function of initial ability level. *Organizational Behavior and Human Performance*, 8, 201–216.
- Laughlin, P., Branch, L., & Johnson, H. (1969). Individual versus triadic performance on a unidimensional complementary task as a function of initial ability level. *Journal of Personality and Social Psychology*, 12(2), 144–150.
- Laughlin, P. & Jaccard, J. (1975). Social facilitation and observational learning of individuals and cooperative pairs. *Journal of Personality and Social Psychology*, 32(5), 873–879.
- Laughlin, P. & Johnson, H. (1966). Group and individual performance on a complementary task as a function of initial ability level. *Journal of Experimental Social Psychology*, 2(4), 407–414.

- Laughlin, P., Kalowski, C., Metzler, M., Ostad, K., & Venclovas, S. (1968). Concept identification as a function of sensory modality, information, and number of persons. *Journal of Experimental Psychology*, 77(2), 335–340.
- Laughlin, P. & McGlynn, R. (1967). Cooperative versus competitive concept attainment as a function of sex and stimulus display. *Journal of Personality and Social Psychology*, 7(4), 398–402.
- Laughlin, P. & McGlynn, R. (1986). Collective induction: Mutual group and individual influence by exchange of hypotheses and evidence. *Journal of Experimental Social Psychology*, 22(6), 567–589.
- Laughlin, P., McGlynn, R., Anderson, J. & Jacobsen, E. (1968). Concept attainment by individuals versus cooperative pairs as a function of memory, sex, and concept rule. *Journal of Personality and Social Psychology*, 8(4), 410–417.
- Lemke, E., Randle, K. & Robertshaw, C. (1969). Effects of degree of initial acquisition, group size, and general mental ability on concept learning and transfer. *Journal of Educational Psychology*, 60(1), 75–78.
- Lorge, I. & Solomon, H. (1960). Group and individual performance in problem solving related to previous exposure to problem, level of aspiration, and group size. *Behavioral Science*, 5, 28–38.
- Lovelace, T. & McKnight, C. (1980). The effects of reading instruction on calculus students' problem solving. *Journal of Reading*, 23(4), 305–308.
- Lynch, B. (1984). Cooperative learning in interdisciplinary education for the allied health professions. *Journal of Allied Health*, 13(2), 83–93.
- McClintock, E. & Sonquist, J. (1976). Cooperative task-oriented groups in a college classroom: A field application. *Journal of Educational Psychology*, 107, 177–183.
- McCurdy, H. & Lambert, W. (1952). The efficiency of small groups in the solution of problems requiring genuine co-operation. *Journal of Personality*, 20, 478–494.
- McDonald, B., Larson, C., Dansereau, D., & Spurlin, J. (1985a). Cooperative dyads: Impact on text learning and transfer. *Contemporary Educational Psychology*, 10, 369–377.
- McGlynn, R., Gibbs, M., & Roberts, S. (1982). Effects of cooperative versus competitive set and coaction on creative responding. *The Journal of Social Psychology*, 118(2), 281–282.
- Michaels, J. (1978). Effects of differential rewarding and sex on math performance. *Journal of Educational Psychology*, 70, 565–573.
- Miller, L. & Hamblin, R. (1963). Interdependence, differential rewarding, and productivity. *American Sociological Review*, 28(5), 768–778.
- Milton, G. (1965). Enthusiasm versus effectiveness in group and individual problem-solving. *Psychological Reports*, 16, 1197–1201.
- Niehoff, B. & Mesch, D. (agosto de 1988). *Effects of reward structures on academic performance and group process behavior in a business school setting*. Trabajo presentado en The Annual Meeting of the Academy of Management, Anaheim, CA.
- Nowicki, S., Duke, M. & Crouch, M. (1978a). Sex differences in locus of control and performance under competitive and cooperative conditions. *Journal of Educational Psychology*, 70(4), 482–486.
- O'Connell, E. (1965). The effect of cooperative and competitive set on the learning of imitation and nonimitation. *Journal of Experimental Social Psychology*, 1, 172–283.
- O'Donnell, A., Dansereau, D., Hall, R., Rocklin, T. (1987). Cognitive, social/affective, and metacognitive outcomes of scripted cooperative learning. *Journal of Educational Psychology*, 79(4), 421–437.
- O'Donnell, A., Dansereau, D., Hythecjer, V., Hall, R., Skaggs, L., Lambiotte, J. & Young, M. (1988). Cooperative procedural learning: Effects of prompting and pre- versus distributed planning activities. *Journal of Educational Psychology*, 80(2), 167–171.
- O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. & Young, M. (1986). The effects of monitoring on cooperative learning. *Journal of Experimental Education*, 54(3), 169–173.
- O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J. & Young, M. (1985). Effects of elaboration frequency on cooperative learning. *Journal of Educational Psychology*, 77(5), 572–580.
- O'Donnell, A., Dansereau, D., Rocklin, T., Lambiotte, J., Hythecker, V. & Larson, C. (1985). Cooperative writing: Direct effects and transfer. *Written Communication*, 2(3), 307–315.
- O'Donnell, A., Dansereau, D., Rocklin, T., Larson, C., Hythecker, V., Young, M., & Lambiotte, J. (1987). Effects of cooperative and individual rewriting on an instruction writing task. *Written Communication*, 4(1), 90–99.
- O'Donnell, A., Larson, C., Dansereau, D. & Rocklin, T. (1986). Effects of cooperation and editing instruction writing performance. *Journal of Experimental Education*, 54(4), 207–210.

- Okun, M. & DiVesta, F. (1975). Cooperation and competition in coacting groups. *Journal of Personality and Social Psychology*, 31(4), 615–620.
- Olson, P. & Davis, J. (1964). Divisible tasks and pooling performance in groups. *Psychological Reports*, 15, 511–517.
- Pawlicki, R. & Gunn, W. (1967). Individual and group performance. *Psychological Reports*, 21, 341–344.
- Perlmutter, H. & De Montmollin, G. (1952). Group learning of nonsense syllables. *Journal of Abnormal and Social Psychology*, 47, 762–769.
- Rabbie, J., Benoist, F., Oosterbann, H., & Visser, L. (1974). Differential power and effects of expected competitive and cooperative intergroup interaction on intragroup and outgroup attitudes. *Journal of Personality and Social Psychology*, 30(1), 404A–A.
- Rabbie, J. & Brey, J. (1971). The anticipation of intergroup cooperation and competition under private and public conditions. *International Journal of Group Tensions*, 1(3), 230–251.
- Rattigan, P. (1985). *The effect of interpersonal cooperation and competition on physical skill, achievement, cognitive achievement and attitudes in volleyball*, (Trabajo final de master), University of Minnesota, Minneapolis.
- Raven, B. & Eachus, H. (1963). Cooperation and competition in means–interdependent triads. *Journal of Abnormal and Social Psychology*, 67(4), 307–316.
- Rewey, K., Dansereau, D., Skaggs, L., Hall, R. & Pitre, U. (1989). Effects of scripted cooperation and knowledge maps on the processing of technical material. *Journal of Educational Psychology*, 81(4), 604–609.
- Roberts, G. (1972). Effect of achievement motivation and social environment on performance of a motor task. *Journal of Motor Behavior*, 4(1), 37–46.
- Rocklin, T., O'Donnell, A., Dansereau, D., Lambiotte, J., Hythecker, V., & Larson, C. (1985). Training learning strategies with computer–aided cooperative learning. *Computers and Education*, 9(1), 67–71.
- Roon, R., Van Pilsum, J., Harris, I., Rosenberg, P., Johnson, R., Liaw, C., & Rosenthal, L. (1983). The experimental use of cooperative learning in a biochemistry class for first year medical students. *Biochemical Education*, 11(1), 12–16.
- Rorie, V. (1979). *The effects of cooperative versus competitive goal structure on student achievement among Afro-American Students*. Manuscrito no publicado, University of Minnesota, Minneapolis.
- Rosenbaum, M., Moore, D., Cotton, J., Cook, M., Hieser, R., Shovar, M. & Gray, M. (1980b). Group productivity and process: Pure and mixed reward structures and task interdependence. *Journal of Personality and Social Psychology*, 39(4), 626–642.
- Rosenfield, D. & Roberts, B. (1979). *The effects of success and failure upon liking for competent and incompetent member of cooperative and competitive groups*. Manuscrito no publicado, Southern Methodist University, Dallas, Texas.
- Rosenfield, D., Stephan, W. G. & Luckner, G. W. (1981). Attraction to competent and incompetent members of cooperative and competitive groups. *Journal of Applied Social Psychology*, 11(4), 416–433.
- Ryack, B. (1965). A comparison of individual and group learning of nonsense syllables. *Journal of Personality and Social Psychology*, 2(2), 296–299.
- Ryen, A. & Kahn, A. (1975). Effects of intergroup orientation on group attitudes and proxemic behavior. *Journal of Personality and Social Psychology*, 31(2), 302–310.
- Schmidt, H. (1982a). Activation and restructuring of prior knowledge and their effects on text processing: Recognition and transfer. In A. Flammar & W. Kintsch (Eds.), *Discourse Processing* (pp.70–84). Amsterdam: North Holland.
- Scott, W. & Cherrington, D. (1974). Effects of competitive, cooperative, and individualistic reinforcement contingencies. *Journal of Personality and Social Psychology*, 30(6), 748–758.
- Seta, J., Paulus, P., & Schkade, J. (1976a). Effects of group size and proximity under cooperative and competitive conditions. *Journal of Personality and Social Psychology*, 34(1), 47–53.
- Shaw, M. (1932). A comparison of individuals and small groups in the rational solution of complex problems. *American Journal of Psychology*, 44, 491–504.
- Shaw, M. (1958a). Some motivational factors in cooperation and competition. *Journal of Personality*, 26, 155–169.
- Sherman, L. (1986). Cooperative vs competitive educational psychology classrooms: A comparative study. *Teaching & Teacher Education*, 2(3), 283–295.
- Silverthorne, C., Chelune, G., & Imada, A. (1974). The effects of competition and cooperation on level of prejudice. *Journal of Social Psychology*, 92, 293–301.



- Sims, V. (1928a). The relative influence of two types of motivation on improvement. *Journal of Educational Psychology*, 19, 480–484.
- Skaggs, L., Dansereau, D., & Hall, R. (1988). *Effects of Cooperative Script Manipulations on Initial Learning and Transfer*. Trabajo presentado en The Annual Meeting of the Southwestern Psychological Association, Tulsa, OK.
- Smith, A., Madden, H. & Sobol, R. (1957). Productivity and recall in cooperative and competitive discussion groups. *Journal of Psychology*, 43, 193–204.
- Smith, K. (1972). The effect of varying reward systems on cooperative game behavior. *Journal of Psychology*, 80, 29–35.
- Smith, M. (1984). *A comparison of cooperative and individualistic learning in associate degree nursing students*. (Tesis doctoral), University of Minnesota.
- Spurlin, J., Dansereau, D., Larson, C., & Brooks, L. (1984). Cooperative learning strategies in processing descriptive text: Effects of role and activity level of the learner. *Cognition and Instruction*, 1(4), 451–463.
- Stephan, C., Burnam, M., & Aronson, E. (1979). Attributions for success and failure after cooperation, competition, or team competition. *European Journal of Social Psychology*, 9, 109–114.
- Stephan, C., Presser, N., Kennedy, J., & Aronson, E. (1978). Attributions to success and failure after cooperative or competitive interaction. *European Journal of Social psychology*, 8, 269–274.
- Taylor, D., Berry, P., & Block, C. (1938). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly*, 3, 23–47.
- Taylor, D. & Faust, W. (1952). Twenty questions: Efficiency in problem solving as a function of size of group. *Journal of Experimental Psychology*, 44, 360–368.
- Thorndike, R. (1938). On what type of task will a group do well? *Journal of Abnormal and Social Psychology*, 33, 409–413.
- Tjosvold, D. (septiembre de 1979). *The effects of unequal power within cooperative and competitive contexts*. Trabajo presentado en The annual meeting of the American Psychological Association, New York.
- Tjosvold, D. (1982a). Effects of cooperative and competitive interdependence and task complexity on subordinates' productivity, perception of leader, and group development. *Canadian Journal of Behavioural Sciences*, 14(1), 24–34.
- Tjosvold, D. (1982b). Effects of approach to controversy on superiors' incorporation of subordinates' information in decision making. *Journal of Applied Psychology*, 67(2), 189–193.
- Tjosvold, D. (1985a). Power and Social Context in Superior-Subordinate Interaction. *Organizational Behavior and Human Decision Processes*, 35, 281–293.
- Tjosvold, D. (1985b). Dynamic within participation: An experimental investigation. *Group and Organization Studies*, 10(3), 260–277.
- Tjosvold, D. (1985c). The effects of attribution and social context in superiors' influence and interaction with low performing subordinates. *Personnel Psychology*, 38(2), 361–376.
- Tjosvold, D. (1988). Effects of shared responsibility and goal interdependence on controversy and decision making between departments. *Journal of Social Psychology*, 128(1), 7–18.
- Tjosvold, D., & Deemer, D. (1980). Effects of controversy within a cooperative or competitive context on organizational decision making. *Journal of Applied Psychology*, 65(5), 590–595.
- Tjosvold, D., & Deemer, D. (1981). Effects of a control or collaborative orientation on participation in organizational decision making. *Canadian Journal of Behavioural Sciences*, 13(1), 33–43.
- Tjosvold, D., & Fabrey, L. (1980). Motivation for perspective taking: Effects of interdependence and dependence on interest in leaning others' intentions. *Psychological Reports*, 46, 755–765.
- Tjosvold, D. & Field, R. (1983). Effects of social context on consensus and majority vote decision making. *Academy of Management Journal*, 26(3), 500–506.
- Tjosvold, D., & Johnson, D. (1978). Controversy within a cooperative or competitive context and cognitive perspective-taking. *Contemporary Educational Psychology*, 3, 376–386.
- Tjosvold, D., Johnson, D. & Johnson, R. (1984). Influence strategy, perspective-taking, and relationships between high- and low-power individuals in cooperative and competitive contexts. *The Journal of Psychology*, 116, 187–202.
- Tjosvold, D., Johnson, D. & Lerner, J. (1981). Effect of affirmation and acceptance on incorporation of opposing information in problem solving. *Journal of Social Psychology*, 114, 103–110.
- Tjosvold, D. & Okun, M. (1979) Effects of unequal power on cooperation in conflict. *Psychological Reports*, 44, 239–242.
- Torrance, E. (1970a). Influence of dyadic interaction on creative functioning. *Psychological Reports*, 26, 391–394.

- Torrance, E. (1971). Stimulation enjoyment and originality in dyadic creativity. *Journal of Educational Psychology*, 62(1), 45–48.
- Turner, V., Alders, C., Hatfield, F., Croy, H. & Sigrist, C. (1966). A study of ways of handling large classes in freshman mathematics. *American Mathematical Monthly*, 73(7), 768–770.
- Van Tuinen, M. & McNeel, S. (1975). A test of the social facilitation theories of Cottrell and Zajonc in coaction situation. *Personality and Social Psychology Bulletin*, 1, 604–607.
- Vaught, G., & Newman, S. (1966). The effects of anxiety on motor–steadiness in competitive and noncompetitive conditions. *Psychonomic Science*, 6(12), 519–520.
- Vroom, V., Grant, L. & Cotton, T. (1969). The consequences of social interaction in group problem solving. *Organizational Behavior and Human Performance*, 4, 77–95.
- Watson, G. (1928). Do groups think more efficiently than individuals? *Journal of Abnormal and Social Psychology*, 23, 328–336.
- Wegner, N. & Zeaman, D. (1956). Team and individual performances on a motor learning task. *The Journal of General Psychology*, 55, 127–142.
- Weinstein, A. & Holzbach, R. (1972). Effects of financial inducement on performance under two task structures. *Proceedings of the 80<sup>th</sup> Annual Conventions of the American Psychological Association*, 7(1), 217–218.
- Wiest, W., Porter, L. & Ghiseli, E. (1961). Relations between individual proficiency and team performance and efficiency. *Journal of Applied Psychology*, 45(6), 435–440.
- Wilson, W. & Miller, N. (1961). Shifts in evaluations of participants following intergroup competition. *Journal of Abnormal and Social Psychology*, 11(2), 110–117.
- Witte, P. (1972). *The effects of group reward structure on interracial acceptance, peer tutoring and academic performance*, (Tesis doctoral), Washington University.
- Worchel, S., Andreoli, V. & Folgeer, R. (1977). Intergroup cooperation and intergroup attraction: The effect of previous interaction and outcome of combined effort. *Journal of Experimental Social Psychology*, 13, 131–140.
- Worchel, S., Axson, D., Ferris, F., Samaha, G. & Schweizer, S. (1978). Determinants of the effect of intergroup cooperation on intergroup attraction. *Journal of Conflict Resolution*, 22(3), 429–438.
- Yuker, H. E. (1955). Group atmosphere and memory. *Journal of Abnormal and Social Psychology*, 51, 17–23.

**Anexo 7. Clasificación Internacional Normalizada de la Educación (CINE)**

<b>Programas generales</b>	<p><b>Programas básicos</b> Programas básicos de educación preescolar, elemental, primaria, secundaria, etc.</p> <p><b>Programas de alfabetización y de aritmética</b> Alfabetización simple y funcional; aritmética elemental.</p> <p><b>Desarrollo personal</b> Desarrollo de destrezas personales, por ejemplo, capacidad de comportamiento, aptitudes intelectuales, capacidad organizativa, programas de orientación.</p>
<b>Educación</b>	<p><b>Formación de personal docente</b> para: educación preescolar, jardines de infancia, escuelas elementales, asignaturas profesionales, prácticas y no profesionales, educación de adultos, formación de personal docente, formación de maestros de niños minusválidos. Programas generales y especializados de formación de personal docente.</p> <p><b>Ciencias de la Educación:</b> elaboración de programas de estudio de materias no profesionales y profesionales. Evaluación de conocimientos, pruebas y mediciones, investigaciones sobre educación; otros programas relacionados con las ciencias de la educación.</p>
<b>Humanidades y Artes</b>	<p><b>Artes</b> Bellas artes: dibujo, pintura y escultura; Artes del espectáculo: música, arte dramático, danza, circo; Artes gráficas y audiovisuales: fotografía, cinematografía, producción musical, producción de radio y televisión, impresión y publicación; Diseño; artesanía.</p> <p><b>Humanidades</b> Religión y teología; Lenguas y culturas extranjeras: lenguas vivas o muertas y sus respectivas literaturas, estudios regionales interdisciplinarios; Lenguas autóctonas: lenguas corrientes o vernáculos y su literatura Otros programas de humanidades: interpretación y traducción, lingüística, literatura comparada, historia, arqueología, filosofía, ética.</p>
<b>Ciencias Sociales, Educación Comercial y Derecho</b>	<p><b>Ciencias Sociales y del comportamiento</b> Economía, historia de la economía, ciencias políticas, sociología, demografía, antropología (excepto antropología física), etnología, futurología, psicología, geografía (excepto geografía física), estudios sobre paz y conflictos, derechos humanos.</p> <p><b>Periodismo e información</b> Periodismo, bibliotecología y personal técnico de bibliotecas, personal técnico de museos y establecimientos similares; Técnicas de documentación; Archivología.</p> <p><b>Educación Comercial y Administración</b> Comercio al por menor, comercialización, ventas, relaciones públicas, asuntos inmobiliarios; Gestión financiera, administración bancaria, seguros, análisis de inversiones; Contabilidad, auditoría, teneduría de libros; Gestión, administración pública, administración institucional, administración de personal; Secretariado y trabajo de oficina.</p> <p><b>Derecho</b> Magistrados locales, notarios, derecho (general, internacional, laboral, marítimo, etc.), jurisprudencia, historia del derecho.</p>
<b>Ciencias</b>	<p><b>Ciencias de la vida</b> Biología, botánica, bacteriología, toxicología, microbiología, zoología, entomología, ornitología, genética, bioquímica, biofísica, otras ciencias afines, excepto medicina y veterinaria.</p> <p><b>Ciencias Físicas</b> Astronomía y ciencias espaciales, física y asignaturas afines, química y asignaturas afines, geología, geofísica, mineralogía, antropología física, geografía física y demás ciencias de la tierra, meteorología y demás ciencias de la atmósfera, comprendida la investigación sobre el clima, las ciencias marinas, vulcanología, paleoecología.</p> <p><b>Matemáticas y Estadística</b> Matemáticas, investigación de operaciones, análisis numérico, ciencias actuariales, estadística y otros sectores afines.</p> <p><b>Informática</b> Informática: Concepción de sistemas, programación informática, procesamiento de datos, redes, sistemas operativos – elaboración de programas informáticos solamente (el material y equipo se deben clasificar en el sector de la Ingeniería).</p>

<b>Ingeniería, Industria y Construcción</b>	<p><b>Ingeniería y profesiones afines</b> Dibujo técnico, mecánica, metalistería, electricidad, electrónica, telecomunicaciones, Ingeniería energética y química, mantenimiento de vehículos, topografía.</p> <p><b>Industria y producción</b> Alimentación y bebidas, textiles, confección, calzado, cuero, materiales (madera, papel, plástico, vidrio, etc.), minería e industrias extractivas.</p> <p><b>Arquitectura y construcción</b> Arquitectura y urbanismo: arquitectura estructural, arquitectura paisajística, planificación comunitaria, cartografía; Edificación, construcción; Ingeniería civil.</p>
<b>Agricultura</b>	<p><b>Agricultura, silvicultura y pesca</b> Agricultura, producción agropecuaria, agronomía, ganadería, horticultura y jardinería, silvicultura y técnicas forestales, parques naturales, flora y fauna, pesca, ciencia y tecnología pesqueras.</p> <p><b>Veterinaria</b> Veterinaria, auxiliar de veterinaria.</p>
<b>Salud y Servicios Sociales</b>	<p><b>Medicina</b> Medicina: anatomía, epidemiología, citología, fisiología, inmunología e inmunohematología, patología, anestesiología, pediatría, obstetricia y ginecología, medicina interna, cirugía, neurología, psiquiatría, radiología, oftalmología; servicios médicos: servicios de salud pública, higiene, farmacia, farmacología, terapéutica, rehabilitación, prótesis, optometría, nutrición; enfermería: enfermería básica, partería; servicios dentales: auxiliar de odontología, higienista dental, técnico de laboratorio dental, odontología.</p> <p><b>Servicios sociales</b> Asistencia social: asistencia a minusválidos, asistencia a la infancia, servicios para jóvenes, servicios de gerontología; Trabajo Social: orientación, asistencia social no clasificados en otra parte</p>
<b>Servicios</b>	<p><b>Servicios personales</b> Hotelería y restaurantes, viajes y turismo, deportes y actividades recreativas, peluquería, tratamientos de belleza y otros servicios personales: lavandería y tintorería, servicios cosméticos, ciencias del hogar.</p> <p><b>Servicios de transporte</b> Formación de marinos, oficiales de marina, náutica, tripulación de aviones, control del tráfico aéreo, transporte ferroviario, transporte por carretera, servicios postales.</p> <p><b>Protección del medio ambiente</b> Conservación, vigilancia y protección del medio ambiente, control de la contaminación atmosférica y del agua, ergonomía y seguridad.</p> <p><b>Servicios de seguridad</b> Protección de personas y bienes: servicios de policía y orden público, criminología, prevención y extinción de incendios, seguridad civil; Educación militar.</p>
<b>Sectores desconocidos o no especificados</b>	(Esta categoría no forma parte de la clasificación en sí, pero en la recopilación de datos “99” se necesita para “los sectores de educación desconocidos o no especificados”.)

## Anexo 8. Manual de codificación

### Variables extrínsecas

1. Número del estudio: identificación de cada estudio con un número para su fácil registro y ubicación en la hoja de Excel.

2. Autor(es): nombre de todos los autores de la publicación respetando el orden en que ellos mismos se han asignado. Se escribe primero el apellido del autor y luego la inicial del nombre. Este proceso se repetirá tantas veces como autores hayan participado en el estudio. Eliminar de los nombres cualquier símbolo que no permita la búsqueda en Excel. Los nombres completos de los autores se separan con comas.

3. Año de publicación: escribir el año en que ha sido publicado el estudio. Si no aparece información se escribe NA.

4. Lustró: período de cinco años.

	Se le asignará el valor
80–84	6
85–89	7
90–94	8
95–99	9
00–04	10
05–09	11
10–12	12
No aparece información	13

5. Década: período de diez años.

	Se le asignará el valor
1980–1989	4
1990–1999	5
2000–2009	6
2010–2012 (por ser este último año el límite superior de tiempo para la búsqueda de estudios)	7
No aparece información	8

6. Título del estudio: escribir el nombre completo del título del estudio.

7. Información del documento: escribir la información necesaria que permita identificar el documento de donde proviene el estudio. Por ejemplo, si se trata de un artículo de investigación se escribirá el nombre completo de la revista, volumen, número y páginas.

8. Tipos de documento: si el estudio ha sido publicado o no. En caso de ser publicado los tipos de documento pueden ser un reporte técnico, disertación, publicaciones profesionales o artículos en revistas científicas. Lo no publicado refiere a que por diversas razones los autores han guardado el estudio

		Se le asignará el valor
Publicados	Reporte técnico (convenciones, conferencias...)	1
	Disertaciones (tesis de maestría o doctorales)	2
	Publicaciones profesionales (libros)	3
	Publicaciones (artículos revistas científicas)	4
No publicados	Investigaciones que tienen los investigadores guardados en sus escritorios	5
No aparece información		6

9. Tipo de fuente: si el estudio proviene de una fuente primaria o secundaria.

		Se asignará el valor
Fuente primaria		3
Fuente secundaria	Meta-análisis	1
	Revisión narrativa	2
	Otro estudio	4

10. Lugar/país: escribir el país de procedencia de las instituciones a cargo del estudio.

11. Lugar/continente: identificar el continente de procedencia de las instituciones a cargo del estudio.

	Se le asignará el valor
Asia	1
Europa	2
África	3
América	5
Oceanía	6
No aparece información	7
Combinado (por ejemplo: 1,2 quiere decir que el estudio ha sido publicado por dos instituciones de lugares de procedencia diferentes, una en Asia y otra en Europa. Los números deben ir separados con comas)	

#### **Variables ecológicas o sustantivas**

12. Asignatura: identificar la disciplina o campo del saber en donde se ha llevado a cabo la intervención.

	Se asignará el valor
Administración	1
Biología	3
Lingüística	4
Economía	5
Educación	6
Enfermería	7
Estadística	8
Física	9
Geografía	10
Ingeniería	11
Matemáticas	15
Música	16
Psicología	17
Química	18
Tecnología	19
No aparece información	20
Otros	21

13. Áreas de conocimiento: agrupación de la disciplina o campo del saber de la asignatura en donde se ha llevado a cabo la intervención según la ciencia de procedencia bajo los criterios establecidos por CINE (Clasificación Internacional Normalizada de la Educación).

	Se asignará el valor
Ciencias	1
Ciencias Sociales, Educación Comercial y Derecho	2
Ingeniería, Industria y Construcción	3
Humanidades y Artes	4
Salud y Servicios Sociales	5
Educación	6

Agricultura	7
Servicios	8
Sectores desconocidos o no especificados	9

14. Conformación de los grupos: los grupos pueden ser homogéneos o heterogéneos según el rendimiento académico y/o sexo.

	Se asignará el valor
Homogéneos	1
Heterogéneos	2
No aparece información	3

15. Número integrantes del grupo: identificar el número de integrantes en cada grupo.

	Se asignará el valor
Parejas	1
Grupo de 2-3	2
Grupos de 3	3
Grupos de 4	4
Grupos de 3-4	5
Grupos de 4-5	6
Grupos de 5	7
Grupos de 5-6	8
Grupos de 6	9
Grupos de 6-7	10
Grupos de 3-4-5	11
Grupos de 4-5-6	12
Grupos de 2, 4	13
Grupos de 2-3-4	14
Grupos de 2, 5	15
Grupos de 4, 6	16
Grupos de 6-7-8	17
Grupos de 8-9-10	18
No aparece información	19

15. Grupos: pares o impares

	Se asignará el valor
Grupos pares	1
Grupos impares	2
No aparece información	3
No aplica	4

16. Número integrantes del grupo: identificar el número de integrantes en cada grupo.

	Se asignará el valor
Dos miembros	(se ha utilizado los valores asignados en el indicador n° 15)
Tres miembros	
Entre cuatro y seis miembros (ambos inclusive)	
No aparece información	

17. Duración de la intervención: período de tiempo en que se ha llevado a cabo la intervención.

	Se asignará el valor
Sesiones (Muy baja)	7
Entre 1 y 4 semanas inclusive (Baja)	1
Entre 5 y 8 semanas (Media)	2
Entre 9 y 12 semanas (Media-Alta)	3

Entre 13 y 16 semanas (Alta)	4
17 semanas o más (Muy alta)	5
No aparece información	6

18. Métodos de aprendizaje cooperativo: identificar todas las técnicas cooperativas que se han utilizado en la intervención para luego analizar solamente aquellos que cumplen con las condiciones de un método cooperativo y dejar fuera los pseudométodos

	Se asignará el valor
Learning Together and Alone (LT)	1
Group Investigation (GI)	3
Jigsaw I	4
Student Teams–Achievement Division (STAD)	6
Team–Assisted Individualization (TAI)	7
Peer Monitoring	11
Collaborative Learning	12
Computer Support (face to face)	13
Cooperative Learning	14
Peer Tutoring	15
Ask Together–Learn Together	20
Jigsaw II	17
Informal Methods	18
Jigsaw IV	19
Subject Jigsaw Technique	21
Small Group	25
Structure Active Learning	26
Cooperative Base Groups	27
Reciprocal Peer Tutoring (RPT)	30
Peer Teaching	31
Cooperative Learning Dyads	35
Peer Collaborative Learning	32
Scripted Cooperation	33
Computer Support (a distance)	37
Combinados	38
Think–Pair–Share	39

19. Clasificación métodos de aprendizaje cooperativo: agrupar los distintos métodos de aprendizaje cooperativo según la siguiente propuesta

<b>STUDENT TEAM LEARNING</b>	Se asignará el valor
STAD	1
TAI	
Jigsaw II	
TAI	
CIRC	
<b>COOPERATIVE INVESTIGATION</b>	Se asignará el valor
Group Investigation (GI)	2
Jigsaw I	
CO-OP CO-OP	
<b>DYADIC METHODS</b>	Se asignará el valor
Scripted Cooperation	3
Reciprocal Peer Tutoring	
Peer Teaching	
Think–Pair–Share	
<b>LEARNING TOGETHER</b>	Se asignará el valor
	4
<b>MÉTODOS COMBINADOS DE AC</b>	Se asignará el valor
STL + PBL	9



Informal Cooperative Learning + Cooperative Learning + Cooperative Base Groups	
Jigsaw I + Ask Together–Learn Together	
STAD + Web-Based Collaborative Learning Instruction (face to face)	
Reciprocal Peer Tutoring + Educational Technology Course (face to face)	
<b>COOPERATIVE LEARNING AND COMPUTER</b>	Se asignará el valor
Cara a cara	11
A distancia	
<b>INFORMAL METHODS</b>	Se asignará el valor
Spontaneous Group Discussion	12
Numbered Heads Together	
Team Product	
Cooperative Review	

20. Estructura de recompensa: en función del método cooperativo, si la recompensa del grupo proviene del rendimiento individual, rendimiento grupal o no hay recompensa

	Se asignará el valor
Recompensas de grupo basadas en los aprendizaje individuales: STAD, TGT, TAI, CIRC y Jigsaw II	1
Recompensas de grupo basadas en el producto grupal: LT	2
Sin recompensas de grupo: GI y Jigsaw I	3
No aplica	4

21. Estructura de la tarea: en función del método cooperativo, si la interdependencia entre los miembros del grupo es alta o baja

	Se asignará el valor
Con especialización: Jigsaw I, Jigsaw II y GI	1
Sin especialización: STAD, TGT, TAI, CIRC y LT	2
No aplica	3

22. Igualdad de oportunidades en la puntuación: en función del método cooperativo, en qué medida el esfuerzo individual contribuye a la puntuación grupal

	Se asignará el valor
Con igualdad de oportunidades: STAD, TGT, Jigsaw II, TAI y CIRC	1
Sin igualdad de oportunidades: LT, GI y Jigsaw I	2
No aplica	3

### Variables metodológicas

21. Asignación: aleatorización o no de sujetos y/o grupos.

	Se asignará el valor
Aleatorización de sujetos	1
Aleatorización de grupos	2
Emparejamiento ( <i>matching technique</i> )	4
No hay aleatorización	0
Aleatorización sujetos y grupos	3

23. Situación basal: antes de la intervención, control o no de la equivalencia entre el grupo control y experimental en rendimiento académico

	Se asignará el valor
Si	1
No	0
No aparece información	3

23. Efectos del investigador (sesgos): si el investigador es el mismo o no que se encarga de llevar a cabo la intervención en el grupo control y experimental

	Se asignará el valor
Igual investigador	1
Diferente investigador	2
No aparece información	3

24. Instrumentos de medida: fiabilidad o no de los instrumentos de medida

	Se asignará el valor
Los instrumentos son estandarizados/validados	1
Los instrumentos no son estandarizados/validado	2
Combinación de instrumentos estandarizados/validados y no estandarizados/validados	3

25. Calidad metodológica: valoración de la calidad metodológica de los estudios

	Se asignará el valor
Baja	1
Media	2
Alta	3

## **Anexo 9. Versión inicial instrumento para medir la calidad metodológica de los estudios (versión en español)**

### **I. INTERVENCIÓN:** Valora la calidad general de la intervención/tratamiento

	SI	NO
La intervención se sustenta en teoría clara y bien articulada		
Las variables principales del estudio están claramente definidas		
Describe la intervención en cuanto a qué se administra y quién la ha recibido		
Describe cómo la intervención difiere del grupo control		
La intervención se ha llevado a cabo en la universidad o espacios vinculados a la comunidad educativa en donde los propios profesores regulares imparten clases. No se ha llevado a cabo la intervención en aulas especializadas creadas y diseñadas por los investigadores para los propósitos del estudio		
Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural porque el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad		
La intervención puede ser replicada por investigadores quienes tienen distintos niveles de formación		

### **II. CONFORMACIÓN DE GRUPOS:** Valora la validez interna del diseño utilizado

	SI	NO
Contempla al menos un grupo experimental y un grupo control		
Hace referencia a si el grupo control no está conformado por sujetos que tenían la opción de participar en la intervención pero se han negado		
El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención		
Se demuestra que no hay diferencias sistemáticas entre el grupo de intervención y de control antes de la intervención (situación basal). Ambos grupos poseen un alto grado de homogeneidad antes de la intervención en cuanto a: niveles de rendimiento académico, datos demográficos...		
El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo		
Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental, es decir, no hay evidencias de contaminación ( <i>cross-over</i> )		

### **III. ASIGNACIÓN:** Valora si las secciones/grupos/sujetos han sido asignados aleatoriamente o si, en caso contrario, se ha controlado el influjo de posibles variables que afecten los resultados

	SI	NO
Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental		
Los grupos han sido asignados aleatoriamente como grupo control o experimental		
Los sujetos han sido asignados aleatoriamente al grupo control o experimental		
No se ha usado la aleatorización para la asignación de sujetos a los grupos pero se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación)		
Advierte sobre cualquier factor que pudiera comprometer el proceso de aleatorización		

### **IV. MUESTRA:** Valora si las secciones/grupos/sujetos han sido asignados aleatoriamente o si, en caso contrario, se han controlado posibles variables que afecten los resultados

	SI	NO
Describe con detalles los criterios de inclusión de la muestra		
Incluye el número de grupos que han sido escogidos del total de la población para luego especificar el tamaño de la muestra		
El estudio incluye un análisis que asegura que el desgaste de la muestra (si lo hay) no socava la equivalencia de los grupos de intervención y de control. El porcentaje de participantes que se ha perdido durante el proceso de recolección de datos es lo suficientemente pequeño que no afecta los resultados del grupo de intervención y de control (no supera el 80% de la muestra original)		
Se ha solicitado a los sujetos de la muestra su consentimiento para participar en la		

investigación antes de conocer si estaban asignados al grupo experimental o control		
---	--	--

**V. RECOLECCIÓN DE DATOS:** Valora los instrumentos y medidas aplicadas

	SI	NO
Los instrumentos utilizados miden con exactitud los resultados reales por los cuales la intervención está destinada a incidir. Existe una correspondencia entre los instrumentos y la intervención realizada		
El grupo control y experimental han sido evaluados con los mismos instrumentos		
Todos los instrumentos aplicados para medir rendimiento han sido validados/estandarizados		
Las medidas de resultado no están a favor del grupo de intervención sobre el grupo control o viceversa. Por ejemplo, un estudio de un programa computarizado para enseñar matemáticas a los jóvenes estudiantes no debe medir los resultados mediante un examen computarizado, ya que el grupo de intervención probablemente tendrá mayor facilidad con el ordenador que el grupo control		

**VI. REPORTE DE RESULTADOS:** Valora los hallazgos encontrados en el estudio

	SI	NO
Describe cómo la intervención ha afectado los resultados		
Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos		
Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos que solo han obtenido un efecto positivo		
Incluye los resultados principales a los cuales responde la pregunta de investigación así como también resultados secundarios de interés		
Reporta los datos de incluso aquellos miembros del grupo de intervención que no participaron o completaron la intervención ( <i>Intention to treat analysis</i> ) como de aquellos sujetos que terminaron la intervención ( <i>Completers analysis</i> )		

**VII. SESGOS DEL INVESTIGADOR:** Valora quien administra la intervención

	SI	NO
El mismo investigador es quien lleva a cabo la intervención en el grupo experimental y control		
Quien aplica la intervención desconoce los objetivos del estudio y en consecuencia a qué grupo pertenece el sujeto que está evaluando (control ciego)		

**VALORACIÓN GLOBAL DE LA CALIDAD DEL ESTUDIO:** Corresponde a la suma de las puntuaciones en cada una de las siete categorías y en la valoración global. Si el indicador ha sido valorado con un NO se puntuará con un valor de 0. Si el indicador ha sido evaluado con un SI se puntuará con un valor de 1. La valoración global de la calidad del estudio oscilará entre 0 pto y 33 pto.

Categorías	Puntuaciones
I. Intervención	/7
II. Conformación de grupos	/6
III. Asignación	/5
IV. Muestra	/4
V. Recolección de datos	/4
VI. Reporte de resultados	/5
VII. Sesgos del investigador	/2
<b>Total calidad global del estudio</b>	<b>/33</b>

**Anexo 10. Versión inicial instrumento para medir la calidad metodológica de los estudios (versión en inglés)**

**I. INTERVENTION:** This evaluates the general intervention/treatment quality

	YES	NO
The intervention is based on a clear and well drafted theory		
The main variables of the study are defined clearly		
It describes the intervention in terms of what is administered and who has received it		
It describes how the intervention differs from the control group		
The intervention has been carried out in the university or in spaces linked to the educational community where the actual regular professors give lectures. The intervention was not carried out in specialised classrooms created and designed by the researchers for the purposes of the study		
Sufficient information is provided to replicate the study in similar situations in a natural environment as the procedure is sufficiently clear, allowing for the execution of the intervention with a high reliability level		
The intervention may be replicated by researchers with different training levels		

**II. GROUP CONFORMATION:** This evaluates the internal validity of the design used

	YES	NO
It considers at least one experimental group and one control group		
It refers to whether the control group is not made up by participants with the option to participate in the intervention but who refused to do so		
The researcher chooses the measurements for result analysis before administering the intervention		
It is proven that there are no systematic differences between the intervention and the control group prior to the intervention (basal location). Both groups are highly homogeneous before the intervention with regard to: academic performance levels, demographic data...		
The study collects the results from the intervention and control groups in the same manner and at the same time		
No participant of the control group has later participated in the experimental group, that is to say, there is no proof of cross-over		

**III. ASSIGNATION:** This evaluates whether the sections/groups/participants were assigned randomly or if, on the contrary, the influence of possible variables affecting the results were controlled

	YES	NO
The classrooms (sections) were assigned randomly as a control or experimental group		
The groups were assigned randomly as a control or experimental group		
The participants were assigned randomly to the control or experimental group		
No randomisation was used for the assignation of participants to the groups, although some strange variable control procedure has been applied (e.g., pairing, blocking or stratification)		
It warns on any factor that may compromise the randomisation process		

**IV. SAMPLE:** This evaluates whether the sections/groups/participants were assigned randomly or if, to the contrary, possible variables affecting the results were controlled

	YES	NO
It describes in detail the sample inclusion criteria		
It includes the number of groups chosen from the population total so as to later specify the sample size		
The study includes an analysis ensuring the sample wear (where applicable) does not hinder the equivalence of the intervention and control groups. The percentage of participants lost during the data collection process is sufficiently small for it not to affect the results of the intervention and control groups (not exceeding 80% of the original sample)		
The sample participants were requested to give their consent to participate in the research before knowing if they were assigned to the experimental or control group		

**V. DATA COLLECTION:** This evaluates the instruments and measurements applied

	YES	NO
The instruments used measure precisely the real results on which the intervention was to have an effect. There is a correspondence between the instruments and the intervention performed		
The control and experimental groups were evaluated using the same instruments		
All performance measurement instruments have been validated/standardised		
The result measurements are not in favour of the intervention group over the control group or vice versa. For example, the study on a computer programme to teach mathematics to young students should not measure its results by means of a computerised examination, as the intervention group will probably find it easier with the computer than the control group		

**VI. REPORTING OF RESULTS:** This evaluates the findings obtained from the study

	YES	NO
It describes how the intervention has affected the results		
It informs on the effects of the intervention in a simple manner for its comprehension using clear and simple terms allowing for the importance of the results obtained to be evaluated		
It informs on the effects of the intervention with regard to all the results the study measured and not only those obtaining a positive effect		
It includes the main results answering the research question, together with secondary results that are of interest		
It reports on the data of even those intervention group members who did not participate in or complete the intervention ( <i>Intention to treat analysis</i> ) as well as that of those participants who completed the intervention ( <i>Completers analysis</i> )		

**VII. RESEARCHER BIAS:** This evaluates who administers the intervention

	YES	NO
It is the same researcher who carries out the intervention on the experimental and control groups		
The person applying the intervention is unaware of the aims of the study and, consequently, to which group the participant he is evaluating (blind control)		

**GLOBAL STUDY QUALITY VALUATION:** This corresponds to the sum of the scores in each of the seven categories and in the global valuation. If the indicator is evaluated with a NO it will receive a score of 0. If the indicator is evaluated with a YES it will receive a score of 1. The global study quality valuation will range between 0 and 33 points.

Categories	Scores
I. Intervention	/7
II. Group Conformation	/6
III. Assignment	/5
IV. Sample	/4
V. Data Collection	/4
VI. Reporting of Results	/5
VII. Researcher Bias	/2
<b>Global Study Quality Total</b>	<b>/33</b>

## **Anexo 11. Solicitud juicio de expertos (versión en español)**

Estimado experto

En estos momentos me encuentro realizando mi tesis doctoral dirigida por el Prof. Dr. Eduardo López López, catedrático de Pedagogía Diferencial en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Universidad Complutense de Madrid, la cual se titula *Efectos del aprendizaje cooperativo e individual en el rendimiento académico en estudiantes de Educación Superior: un meta-análisis*.

Se entiende por meta-análisis el análisis estadístico de una gran colección de resultados de trabajos individuales con el propósito de integrar los hallazgos obtenidos o también como la síntesis estadística de los datos de estudios comparables, que proporciona un resumen numérico de los resultados globales.

El objetivo general de la investigación es sistematizar estadísticamente estudios que midan los efectos del aprendizaje cooperativo en comparación con el aprendizaje individual en el rendimiento académico de estudiantes universitarios y así ofrecer recomendaciones y sugerencias para la mejora de la práctica educativa sustentada en la evidencia empírica.

Es por ello que solicito su colaboración para **evaluar la escala de calidad metodológica de estudios experimentales y cuasiexperimentales** que se ha diseñado con la finalidad de incluir esta característica de los estudios para su posterior comprobación por ser considerada una posible variable moderadora en los resultados.

El instrumento se ha elaborado en función de siete grandes dimensiones: intervención, conformación de grupos, asignación, muestra, recolección de datos, reporte de resultados y sesgos del investigador y es el resultado de la revisión de varios documentos, entre ellos:

- U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003). *Identifying and implementing educational practices supported by rigorous evidence: a user friendly guide*. Recuperado el 29 de abril de 2012 de <http://www2.ed.gov/rschstat/research/pubs/rigoroussevid/index.html>
- Coalition for Evidence-Based Policy (2009). *Recommendation on criteria for establishing strong evidence of effectiveness*. Recuperado el 29 de abril de 2012 de [http://coalition4evidence.org/wordpress/?page\\_id=312](http://coalition4evidence.org/wordpress/?page_id=312)
- Coalition for Evidence-Based Policy (2010). *Checklist for reviewing a randomized controlled trial of a social program or project, to assess whether it produced valid evidence*. Recuperado el 29 de abril de 2012 de <http://coalition4evidence.org/>
- Higgins, J. & Green S. (2011) (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration, 2011. Recuperado de [www.cochrane-handbook.org](http://www.cochrane-handbook.org).

Su participación en el estudio es de carácter anónimo y voluntario.

Contando con su valoración como experto **antes del**

Quedando a su disposición para cualquier duda

Se despide,

Celia Camilli Trujillo  
Universidad Complutense de Madrid

### JUICIO DE EXPERTOS

Doctor en Por la universidad Universidad en donde trabaja actualmente Experto en
---

1. ¿La escala responde al objetivo general del estudio?  
 Si \_\_\_\_ No \_\_\_\_ En caso de ser negativa su respuesta explique el por qué

---



---

2. Los dos criterios de valoración utilizados “SI” o “NO” ¿Son adecuados?  
 Si \_\_\_\_ No \_\_\_\_ En caso de ser negativa su respuesta explique el por qué

---



---

3. Las instrucciones para contestar la escala de valoración ¿Son claras y precisas?  
 Si \_\_\_\_ No \_\_\_\_ En caso de ser negativa su respuesta explique el por qué

---



---

4. La presentación de la escala en cuanto al formato (tipo y tamaño de letra, distribución del contenido...) ¿Es adecuado?  
 Si \_\_\_\_ No \_\_\_\_ En caso de ser negativa su respuesta explique el por qué

---



---

5. Las siete categorías que conforman la escala ¿Son suficientes?  
 Si \_\_\_\_ No \_\_\_\_ En caso de ser negativa su respuesta explique el por qué

---



---

6. Puntúe en una escala del 0 al 2 cada uno de los aspectos en donde 0 representa el menor valor y 2 la máxima puntuación

#### I. INTERVENCIÓN: Valora la calidad general de la intervención/tratamiento

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
La intervención se sustenta en teoría clara y bien articulada					
Las variables principales del estudio están					



claramente definidas					
Describe la intervención en cuanto a qué se administra y quién la ha recibido					
Describe cómo la intervención difiere del grupo control					
La intervención se ha llevado a cabo en la universidad o espacios vinculados a la comunidad educativa en donde los propios profesores regulares imparten clases. No se ha llevado a cabo la intervención en aulas especializadas creadas y diseñadas por los investigadores para los propósitos del estudio					
Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural porque el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad					
La intervención puede ser replicada por investigadores quienes tienen distintos niveles de formación					

## II. CONFORMACIÓN DE GRUPOS: Valora la validez interna del diseño utilizado

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
Contempla al menos un grupo experimental y un grupo control					
Hace referencia a si el grupo control no está conformado por sujetos que tenían la opción de participar en la intervención pero se han negado					
El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención					
Se demuestra que no hay diferencias sistemáticas entre el grupo de intervención y de control antes de la intervención (situación basal). Ambos grupos poseen un alto grado de homogeneidad antes de la intervención en cuanto a: niveles de rendimiento académico, datos demográficos...					
El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo					
Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental, es decir, no hay evidencias de contaminación ( <i>cross-over</i> )					

## III. ASIGNACIÓN: Valora si las secciones/grupos/sujetos han sido asignados aleatoriamente o si, en caso contrario, se ha controlado el influjo de posibles variables que afecten los resultados

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental					
Los grupos han sido asignados aleatoriamente como grupo control o experimental					
Los sujetos han sido asignados aleatoriamente al grupo control o experimental					
No se ha usado la aleatorización para la asignación de sujetos a los grupos pero se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación)					
Advierte sobre cualquier factor que pudiera comprometer el proceso de aleatorización					

**IV. MUESTRA:** Valora si las secciones/grupos/sujetos han sido asignados aleatoriamente o si, en caso contrario, se han controlado posibles variables que afecten los resultados

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
Describe con detalles los criterios de inclusión de la muestra					
Incluye el número de grupos que han sido escogidos del total de la población para luego especificar el tamaño de la muestra					
El estudio incluye un análisis que asegura que el desgaste de la muestra (si lo hay) no socava la equivalencia de los grupos de intervención y de control. El porcentaje de participantes que se ha perdido durante el proceso de recolección de datos es lo suficientemente pequeño que no afecta los resultados del grupo de intervención y de control (no supera el 80% de la muestra original)					
Se ha solicitado a los sujetos de la muestra su consentimiento para participar en la investigación antes de conocer si estaban asignados al grupo experimental o control					

**V. RECOLECCIÓN DE DATOS:** Valora los instrumentos y medidas aplicadas

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
Los instrumentos utilizados miden con exactitud los resultados reales por los cuales la intervención está destinada a incidir. Existe una correspondencia entre los instrumentos y la intervención realizada					
El grupo control y experimental han sido evaluados con los mismos instrumentos					
Todos los instrumentos aplicados para medir					

rendimiento han sido validados/estandarizados					
Las medidas de resultado no están a favor del grupo de intervención sobre el grupo control o viceversa. Por ejemplo, un estudio de un programa computarizado para enseñar matemáticas a los jóvenes estudiantes no deben medir los resultados mediante un examen computarizado, ya que el grupo de intervención probablemente tendrá mayor facilidad con el ordenador que el grupo control					

**VI. REPORTE DE RESULTADOS:** Valora los hallazgos encontrados en el estudio

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
Describe cómo la intervención ha afectado los resultados					
Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos					
Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos que solo han obtenido un efecto positivo					
Incluye los resultados principales a los cuales responde la pregunta de investigación así como también resultados secundarios de interés					
Reporta los datos de incluso aquellos miembros del grupo de intervención que no participaron o completaron la intervención ( <i>Intention to treat analysis</i> ) como de aquellos sujetos que terminaron la intervención ( <i>Completers analysis</i> )					

**VII. SESGOS DEL INVESTIGADOR:** Valora quien administra la intervención

	Redacción	Orden	Contenido	Pertinencia	Observaciones/recomendaciones
El mismo investigador es quien lleva a cabo la intervención en el grupo experimental y control					
Quien aplica la intervención desconoce los objetivos del estudio y en consecuencia a qué grupo pertenece el sujeto que está evaluando (control ciego)					

Otras sugerencias:

---



---



---



---

7. Valoración global de la escala (Siendo el valor 5 la puntuación más alta)  
1 \_\_\_\_ 2 \_\_\_\_ 3 \_\_\_\_ 4 \_\_\_\_ 5 \_\_\_\_

## **Anexo 12. Solicitud juicio de expertos (versión en inglés)**

Dear professor

At present I am preparing my doctoral thesis entitled *Effects of cooperative and individual learning on student academic achievement in higher education: a meta-analysis*, under the direction of Prof. Dr. Eduardo López López, Differential Pedagogy lecturer at the Department of Research Methods and Education Diagnosis in the Complutense University of Madrid.

Meta-analysis is understood like a statistical analysis of a large collection of results from individual pieces of work with the purpose of integrating the findings obtained or also the statistical synthesis of data from comparable studies, providing a numerical summary of the global results.

The general aim of the study is to statistically systemise quality studies linked to the effects of Cooperative Learning as a teaching/learning methodology in the academic performance of university students, with the purpose of offering recommendations and suggestions to improve the educational practice based on empirical evidence.

This is the reason for me to request your collaboration in the **evaluation of the quality scale of experimental and quasi-experimental studies**, designed to include this feature of the studies for further testing to be considered a possible moderating variable in the results.

The purpose of the instrument is the valuation of the quality of **experimental or quasi-experimental design primary studies** with regard to seven large categories: Intervention, Group Conformation, Assignment, Sample, Data Collection, Reporting of Results and, Researcher Bias. The categories assumed in the instrument are the result of reviewing several documents, among them:

- U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003). *Identifying and implementing educational practices supported by rigorous evidence: a user friendly guide*. Recovered on 29<sup>th</sup> April 2012 from <http://www2.ed.gov/rschstat/research/pubs/rigoroussevid/index.html>
- Coalition for Evidence-Based Policy (2009). *Recommendation on criteria for establishing strong evidence of effectiveness*. Recovered on 29<sup>th</sup> April 2012 from [http://coalition4evidence.org/wordpress/?page\\_id=312](http://coalition4evidence.org/wordpress/?page_id=312)
- Coalition for Evidence-Based Policy (2010). *Checklist for reviewing a randomized controlled trial of a social program or project, to assess whether it produced valid evidence*. Recovered on 29<sup>th</sup> April 2012 from <http://coalition4evidence.org/>
- Higgins, J. & Green S. (2011) (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration, 2011. Recuperado de [www.cochrane-handbook.org](http://www.cochrane-handbook.org).

Your participation in the study will be completely anonymous and voluntary.

I look forward to your evaluation as an expert **before**

Yours sincerely,

Celia Camilli Trujillo  
Complutense University of Madrid  
Spain

**EXPERT JUDGEMENT**

Doctor in

By the university of

University in which you work at present

Expert in

1. Does the scale respond to the general aim of the study?

Yes \_\_\_\_ No \_\_\_\_ Explain why in the case of a negative answer

---

---

---

2. Are the “YES” or “NO” valuation criteria appropriate?

Yes \_\_\_\_ No \_\_\_\_ Explain why in the case of a negative answer

---

---

---

3. Are the instructions to answer the valuation criteria clear and precise?

Yes \_\_\_\_ No \_\_\_\_ Explain why in the case of a negative answer

---

---

---

4. Is the presentation appropriate with regard to its format (type and font size, content distribution...)?

Yes \_\_\_\_ No \_\_\_\_ Explain why in the case of a negative answer

---

---

---

5. Are the seven categories making up the scale sufficient?

Yes \_\_\_\_ No \_\_\_\_ Explain why in the case of a negative answer

---

---

---

6. Score each of the aspects on a scale between 0 and 2, where 0 represents the lowest value and 2 the maximum score

**I. INTERVENTION:** This evaluates the general intervention/treatment quality

	Composition	Order	Content	Relevance	Observations/recommendations
The intervention is based on a clear and well drafted theory					
The main variables of the study are defined clearly					
It describes the intervention in terms of what is administered and who has received it					
It describes how the intervention differs from the control group					
The intervention has been carried out in the university or in spaces linked to the educational community where the actual regular professors give lectures. The intervention was not carried out in specialised classrooms created and designed by the researchers for the purposes of the study					
Sufficient information is provided to replicate the study in similar situations in a natural environment as the procedure is sufficiently clear, allowing for the execution of the intervention with a high reliability level					
The intervention may be replicated by researchers with different training levels					

**II. GROUP CONFORMATION:** This evaluates the internal validity of the design used

	Composition	Order	Content	Relevance	Observations/recommendations
It considers at least one experimental group and one control group					
It refers to whether the control group is not made up by participants with the option to participate in the intervention but who refused to do so					
The researcher chooses the measurements for result analysis before administering the intervention					
It is proven that there are no systematic differences between the intervention and the control group prior to the intervention (basal location). Both groups are highly homogeneous before the intervention with regard to: academic performance levels, demographic data...					
The study collects the results from the intervention and control groups in the same manner and at the same time					
No participant of the control group has later					

participated in the experimental group, that is to say, there is no proof of cross-over					
---	--	--	--	--	--

**III. ASSIGNATION:** This evaluates whether the sections/groups/participants were assigned randomly or if, on the contrary, the influence of possible variables affecting the results were controlled

	Composition	Order	Content	Relevance	Observations/recommendations
The classrooms (sections) were assigned randomly as a control or experimental group					
The groups were assigned randomly as a control or experimental group					
The participants were assigned randomly to the control or experimental group					
No randomisation was used for the assignation of participants to the groups, although some strange variable control procedure has been applied (e.g., pairing, blocking or stratification)					
It warns on any factor that may compromise the randomisation process					

**IV. SAMPLE:** This evaluates whether the sections/groups/participants were assigned randomly or if, to the contrary, possible variables affecting the results were controlled

	Composition	Order	Content	Relevance	Observations/recommendations
It describes in detail the sample inclusion criteria					
It includes the number of groups chosen from the population total so as to later specify the sample size					
The study includes an analysis ensuring the sample wear (where applicable) does not hinder the equivalence of the intervention and control groups. The percentage of participants lost during the data collection process is sufficiently small for it not to affect the results of the intervention and control groups (not exceeding 80% of the original sample)					
The sample participants were requested to give their consent to participate in the research before knowing if they were assigned to the experimental or control group					

**V. DATA COLLECTION:** This evaluates the instruments and measurements applied

	Composition	Order	Content	Relevance	Observations/recommendations
The instruments used measure precisely the real results on which the intervention was to have an effect. There is a correspondence between the					



instruments and the intervention performed					
The control and experimental groups were evaluated using the same instruments					
All performance measurement instruments have been validated/standardised					
The result measurements are not in favour of the intervention group over the control group or vice versa. For example, the study on a computer programme to teach mathematics to young students should not measure its results by means of a computerised examination, as the intervention group will probably find it easier with the computer than the control group					

**VI. REPORTING OF RESULTS:** This evaluates the findings obtained from the study

	Composition	Order	Content	Relevance	Observations/recommendations
It describes how the intervention has affected the results					
It informs on the effects of the intervention in a simple manner for its comprehension using clear and simple terms allowing for the importance of the results obtained to be evaluated					
It informs on the effects of the intervention with regard to all the results the study measured and not only those obtaining a positive effect					
It includes the main results answering the research question, together with secondary results that are of interest					
It reports on the data of even those intervention group members who did not participate in or complete the intervention ( <i>Intention to treat analysis</i> ) as well as that of those participants who completed the intervention ( <i>Completers analysis</i> )					

**VII. RESEARCHER BIAS:** This evaluates who administers the intervention

	Composition	Order	Content	Relevance	Observations/recommendations
It is the same researcher who carries out the intervention on the experimental and control groups					
The person applying the intervention is unaware of the aims of the study and, consequently, to which group the participant he is evaluating (blind control)					

Other suggestions:

---



---



---



---

7. Global scale valuation (with 5 being the highest score)

1\_\_\_\_ 2\_\_\_\_ 3\_\_\_\_ 4\_\_\_\_ 5\_\_\_\_

Anexo 13. Resultados juicio de expertos: frecuencias, porcentajes y coeficiente Alfa Cronbach

		Redacción			Orden			Contenido			Pertinencia			Alpha Cronbach
		0	1	2	0	1	2	0	1	2	0	1	2	
1. Intervención	Ítem 1.1		8 (26,7%)	22 (73,3%)		1 (3,6%)	27 (96,4%)		5 (17,9%)	23 (82,1%)		5 (16,7%)	25 (83,3%)	0,87
	Ítem 1.2		4 (13,3%)	26 (86,7%)			28 (100%)		5 (17,9%)	23 (82,1%)		2 (6,7%)	28 (93,3%)	
	Ítem 1.3	2 (6,9%)	8 (27,6%)	19 (65,5%)			27 (100%)		5 (18,5%)	22 (81,5%)		3 (10,3%)	26 (89,7%)	
	Ítem 1.4	1 (3,4%)	12 (41,4%)	16 (55,2%)	1 (3,7%)	1 (3,7%)	25 (92,6%)	1 (3,7%)	4 (14,8%)	22 (81,5%)	1 (3,4%)	5 (17,2%)	23 (79,3%)	
	Ítem 1.5	6 (20%)	14 (46,7%)	10 (33,3%)		2 (7,1%)	26 (92,9%)	4 (14,3%)	7 (25%)	17 (60,7%)	2 (6,7%)	4 (13,3%)	24 (80%)	
	Ítem 1.6	4 (13,3%)	12 (40%)	14 (46,7%)		1 (3,6%)	27 (96,4%)	1 (3,6%)	8 (28,6%)	19 (67,9%)		2 (6,7%)	28 (93,3%)	
	Ítem 1.7	6 (20%)	10 (33,3%)	14 (46,7%)	1 (3,6%)	2 (7,1%)	25 (89,3%)	5 (17,9%)	3 (10,7%)	20 (71,4%)	4 (13,3%)	4 (13,3%)	22 (73,3%)	
2. Conformación de grupos	Ítem 2.1	2 (6,7%)	4 (13,3%)	24 (80%)		1 (3,6%)	27 (96,4%)	2 (7,1%)	2 (7,1%)	24 (85,7%)		1 (3,4%)	28 (96,6%)	0,762
	Ítem 2.2	6 (20,7%)	9 (31%)	14 (48,3%)	1 (3,7%)	3 (11,1%)	23 (85,2%)	5 (18,5%)	7 (25,9%)	15 (55,6%)	1 (3,4%)	8 (27,6%)	20 (69%)	
	Ítem 2.3	3 (10,7%)	3 (10,7%)	22 (78,6%)		2 (7,7%)	24 (92,3%)	1 (3,8%)	6 (23,1%)	19 (73,1%)		4 (14,8%)	23 (85,2%)	
	Ítem 2.4	8 (26,7%)	8 (26,7%)	14 (46,7%)	1 (3,6%)	3 (10,7%)	24 (85,7%)	3 (10,7%)	4 (14,3%)	21 (75%)	1 (3,3%)	1 (3,3%)	28 (93,3%)	
	Ítem 2.5	3 (10%)	3 (10%)	24 (80%)			28 (100%)	3 (10,7%)	3 (10,7%)	22 (78,6%)		2 (6,7%)	28 (93,3%)	
3. Asignación	Ítem 2.6	4 (13,3%)	5 (16,7%)	21 (70%)	1 (3,6%)	1 (3,6%)	26 (92,9%)	6 (21,4%)	2 (7,1%)	20 (71,4%)	1 (3,3%)	3 (10%)	26 (86,7%)	0,801
	Ítem 3.1	2 (6,9%)	5 (17,2%)	22 (75,9%)		2 (7,4%)	25 (92,6%)	1 (3,7%)	3 (11,1%)	23 (85,2%)		4 (13,8%)	25 (86,2%)	
	Ítem 3.2	4 (13,3%)	3 (10%)	23 (76,7%)	1 (3,6%)	1 (3,6%)	26 (92,9%)	2 (7,1%)	2 (7,1%)	24 (85,7%)	2 (6,9%)	2 (6,9%)	25 (86,2%)	
	Ítem 3.3		3 (10%)	27 (90%)	1 (3,6%)		27 (96,4%)		1 (3,6%)	27 (96,4%)	1 (3,3%)	2 (6,7%)	27 (90%)	
	Ítem 3.4	2 (6,7%)	10 (33,3%)	18 (60%)		1 (3,6%)	27 (96,4%)	3 (10,7%)	3 (10,7%)	22 (78,6%)		2 (6,7%)	28 (93,3%)	
4. Muestra	Ítem 3.5	2 (6,9%)	6 (20,7%)	21 (72,4%)		1 (3,7%)	26 (96,3%)	2 (7,4%)	6 (22,2%)	19 (70,4%)	2 (6,9%)	6 (20,7%)	21 (72,4%)	0,786
	Ítem 4.1	1 (3,3%)	3 (10%)	26 (86,7%)		1 (3,6%)	27 (96,4%)		4 (14,3%)	24 (85,7%)		2 (6,7%)	28 (93,3%)	
	Ítem 4.2	3 (11,1%)	5 (18,5%)	19 (70,4%)		1 (4%)	24 (96%)	1 (4%)	2 (8%)	22 (88%)	1 (3,7%)	4 (14,8%)	22 (81,5%)	
	Ítem 4.3	5 (16,7%)	13 (43,3%)	12 (40%)	1 (3,6%)	1 (3,6%)	26 (92,9%)	1 (3,6%)	5 (17,9%)	22 (78,6%)		2 (6,7%)	28 (93,3%)	
	Ítem 4.4	1 (3,4%)	5 (17,2%)	23 (79,3%)		1 (3,7%)	26 (96,3%)	1 (3,7%)	1 (3,7%)	25 (92,6%)	3 (10%)	3 (10%)	24 (80%)	
5. Recolección de datos	Ítem 5.1	3 (10,3%)	12 (41,4%)	14 (48,3%)		2 (7,4%)	25 (92,6%)	1 (3,7%)	5 (18,5%)	21 (77,8%)		4 (13,8%)	25 (86,2%)	0,845
	Ítem 5.2		4 (14,3%)	24 (85,7%)		2 (7,7%)	24 (92,3%)	2 (7,7%)	2 (7,7%)	22 (84,6%)	1 (3,6%)	1 (3,6%)	26 (92,9%)	

	Ítem 5.3	1 (3,4%)	6 (20,7%)	22 (75,9%)		1 (3,7%)	26 (96,3%)	2 (7,4%)	5 (18,5%)	20 (74,1%)		3 (10,3%)	26 (89,7%)	
	Ítem 5.4	3 (10,7%)	7 (25%)	18 (64,3%)	2 (7,7%)	1 (3,8%)	23 (88,5%)	2 (7,7%)		24 (92,3%)	1 (3,6%)	3 (10,7%)	24 (85,7%)	
6. Resultados	Ítem 6.1	2 (6,7%)	5 (16,7%)	23 (76,7%)		1 (3,6%)	27 (96,4%)		4 (14,3%)	24 (85,7%)	1 (3,3%)	2 (6,7%)	27 (90%)	0,784
	Ítem 6.2	1 (3,4%)	4 (13,8%)	24 (82,8%)			27 (100%)	1 (3,7%)	2 (7,4%)	24 (88,9%)	4 (13,3%)	1 (3,3%)	25 (83,3%)	
	Ítem 6.3		8 (26,7%)	22 (73,3%)		1 (3,6%)	27 (96,4%)		2 (7,1%)	26 (92,9%)	1 (3,3%)	3 (10%)	26 (86,7%)	
	Ítem 6.4		7 (24,1%)	22 (75,9%)			27 (100%)		2 (7,4%)	25 (92,6%)	4 (13,8%)	1 (3,4%)	24 (82,8%)	
7. Sesgos del investigador	Ítem 6.5	3 (10,7%)	5 (17,9%)	20 (71,4%)		1 (3,8%)	25 (96,2%)		2 (7,7%)	24 (92,3%)	2 (7,1%)	4 (14,3%)	22 (78,6%)	0,876
	Ítem 7.1		7 (23,3%)	23 (76,7%)		2 (7,1%)	26 (92,9%)		6 (21,4%)	22 (78,6%)	1 (3,3%)	3 (10%)	26 (86,7%)	
	Ítem 7.2	2 (6,9%)	3 (10,3%)	24 (82,8%)	1 (3,7%)		26 (96,3%)	1 (3,7%)	3 (11,1%)	23 (85,2%)	1 (3,4%)	2 (6,9%)	26 (89,7%)	
Alfa Cronbach		0,926			0,864			0,847			0,965			0,941

#### Anexo 14. Comentarios y sugerencias generales de los expertos

COMENTARIOS Y SUGERENCIAS GENERALES
<p>Pregunta 1. ¿La escala responde al objetivo general del estudio?</p> <ul style="list-style-type: none"> <li>– Es muy adecuada y pertinente.</li> <li>– Sólo parcialmente. En primer lugar, creo que deja fuera la idea central de la validez de constructo, desde la cual todo cobra sentido. Por otro lado, creo que quizá sea una buena herramienta para que tu valores la calidad de las investigaciones primarias, pero no sé si vas a tener mucho éxito en que otros académicos te valoren un conjunto de investigaciones. De hecho creo que no debes dejar esa responsabilidad a nadie más. Por último, la calidad de la investigación es algo subjetivo, está en el ojo de quien la lee, por eso creo que debe ser el mismo ojo quien lo valore. Sobre este punto hay mucha discusión en la literatura meta-analítica, pues no deja de ser un sesgo excluir a las investigaciones más débiles metodológicamente.</li> <li>– Hay que mejorarla siguiendo los comentarios que aportó al margen. No estoy segura de que una investigación con valoración de 25, por ejemplo, tuviera mejor calidad que una valoración con 20. ¿Todos los ítems tienen el mismo peso? Habría que justificarlo.</li> <li>– Yes. It does, but as noted below (I didn't review/edit) all, many of the items are ambiguous, overly complex, and/or need further English edition.</li> <li>– If the scale is indeed meant for RCRs, then the scale may be OK. The scale is less applicable to nonexperimental designs, although these can also give meaningful results.</li> <li>– This is a reasonable scale for the reviewer and will provide sufficient data for the analysis.</li> </ul>
<p>Pregunta 2. Los dos criterios de valoración utilizados “SI” o “NO” ¿Son adecuados?</p> <ul style="list-style-type: none"> <li>– Aunque en la mayoría de casos probablemente sean suficientes, puede haber casos en que sea apropiado una opción de respuesta tipo 'No reportado'.</li> <li>– Sí, pero es incompleto es mejor tener en cuenta una escala 1–4 o deseablemente de 1 a 6, ya que este tipo de investigaciones si quiere hacer un tratamiento estadístico a posteriori debe ser cuantificable.</li> <li>– Quizá podría incluirse una categoría intermedia para algunos ítems en los que la respuesta puede no ser tan clara como SI o NO. Por ejemplo el primer ítem: la intervención se sustenta en una teoría clara y bien articulada; o el ítem 6: ofrece suficiente información...</li> <li>– Sería aconsejable incluir una categoría central, y cambiar la escala en función del grado de acuerdo o desacuerdo.</li> <li>– Si es instrumento se aplica a una muestra considerable de sujetos. En caso que sea una muestra reducida sugiero buscar profundizar en algunas preguntas.</li> <li>– Aunque es apropiado el criterio SI o NO, debido al tipo de investigación estadística, en algunos indicadores parece que hace falta alguna explicación adicional, o indicadores muy largos que puede ser difícil que sea señalados como un "SI" si no se cumplen todas las condiciones. En estos casos sugiero fraccionar o mejorar la redacción.</li> <li>– Sí en casi todos los ítems. Pero a veces pareciera ser útil colocar un tercer criterio que señale NR y significaría no es reportado, aunque al no ser reportado se asumiría que no, pero es algo a considerar.</li> <li>– Creo que No es suficiente para reflejar todos los elementos de la escala. De hecho una escala tipo Likert con categorías pares sería mas adecuada para valorar muchos de los aspectos por los que preguntas.</li> <li>– No siempre. El algún ítem, según indico al margen, debería utilizarse una escala Likert.</li> <li>– Pero algunos síes tendrían a mi modo de ver que tener una mayor ponderación (valer más de un punto): aleatorización, grupo experimental y control, medidas fiables, evaluadores ciegos, resultados de las intervenciones convertibles en tamaños del efecto. Otra opción es convertir estos elementos en criterios de inclusión en el meta-análisis y por tanto TODOS los estudios cumplirían y no sería necesario incluirlos en los criterios de valoración.</li> <li>– No, see my coment, please.</li> <li>– The scale below suggests 0–2????</li> <li>– No, I would al least add 'not applicable' and 'not known'. Second, a problem I have with simply summing up the scores, is that in this way each issue receives the same weight. However, some issues seem more crucial to me than others. For instance, if there is no control group, then I would definitely say that the quality is low, even if the total score on the scale is high...</li> <li>– Yes, keep it simple.</li> <li>– Si, sin embargo corregiría la opción del 0 al 2 porque en lugar de sí o no admite la posibilidad de un "1" que añadiría a la medición otra 3era opción. Lo dejaría como "sí" o "no".</li> </ul>
<p>Pregunta 3. Las instrucciones para contestar la escala de valoración ¿Son claras y precisas?</p>

<ul style="list-style-type: none"> <li>– The scoring of the global evaluation scale (the last item in the questionnaire was incorrectly entered in terms of the n. of questions in each dimension).</li> <li>– No hay o no las he sabido encontrar. La única información que he visto para cada ítem es el ítem mismo.</li> <li>– Es clara.</li> <li>– Creo que es algo confuso que en las instrucciones se pide que se haga una "valoración del aprendizaje cooperativo como metodología de enseñanza" cuando en realidad lo que se quiere es que se valore la calidad metodológica del estudio analizado.</li> <li>– "Las investigaciones seleccionadas tienen como objetivo evaluar el efecto que el aprendizaje cooperativo como método de enseñanza produce en el rendimiento académico de estudiantes universitarios. En primer lugar, lea detenidamente las investigaciones que le han sido asignadas. En segundo lugar responda al cuestionario por cada investigación. El cuestionario está compuesto por 7 categorías (especificar categorías). En cada categoría se hacen referencia a una serie de aspectos. Lea detenidamente las cuestiones que se le plantean y señale con una X la opción que más se ajuste a su valoración.</li> <li>– Esta parte es necesaria leerla varias veces y aún así sigue sin quedar claro a qué se refiere "Luego de la lectura de las investigaciones que le han sido asignadas seleccione con una (X) la opción que más se ajuste en la valoración del aprendizaje cooperativo como metodología de enseñanza y su efecto en el rendimiento académico de estudiantes universitarios."</li> <li>– Si en la mayoría de los ítems, en aquellos de dudas, ambivalencia y preguntar más de una cosa en cada ítem está reportado con un comentario en el instrumento muestra enviado.</li> <li>– Las instrucciones si, pero falta contexto del meta-análisis. Específicamente el constructo, como he mencionado antes.</li> <li>– No siempre. Hay cuestiones terminológicas por el uso del español americano y el abuso de las oraciones negativas que dificultan la comprensión. También hay redundancias.</li> <li>– Algunas instrucciones son prolijas, especialmente en aquellas oraciones largas: se pueden simplificar sin merma en cuanto a su contenido. A este respecto, no queda claro si el evaluador realiza su tarea sobre uno o varios aspectos del estudio.</li> <li>– Quizá en el párrafo: "El objetivo general del presente estudio es sistematizar estadísticamente estudios de calidad vinculados..." se debería indicar la naturaleza de los estudios de calidad cuya sistematización se solicita.</li> <li>– Aunque no parece imprescindible la referencia a las fuentes de donde se han tomado los criterios. Suprimiéndole se reduciría la información aportada.</li> <li>– No, Make sure you clarify what to answer when something is not well described, is not reported, or is not clear.</li> <li>– Not always, see comments below.</li> <li>– No. The first sentence is awkwardly worded and provides two different criteria to evaluate no, ver revisión.</li> </ul>
<p>Pregunta 4. La presentación de la escala en cuanto al formato ¿Es adecuado?</p> <ul style="list-style-type: none"> <li>– Yo he leído la escala en el ordenador y tenía puesto el zoom a 140%.</li> <li>– Presenta un buen formato.</li> <li>– Creo que podría ayudar, por lo menos para la valoración del instrumento, el numerar los ítems</li> <li>– Demasiado contenido por cada ítem, se hace largo. Intentaría reducir el contenido.</li> <li>– NO, needs refinement, clarity, clearer operational descriptions of criteria.</li> <li>– The format is good. It is unclear what would be entered in the box "RESEARCH 1".</li> <li>– Si, con los matices y sugerencias que se exponen en la corrección.</li> </ul>
<p>Pregunta 5. Las siete categorías que conforman la escala ¿Son suficientes?</p> <ul style="list-style-type: none"> <li>– Las dimensiones entendemos que son pertinentes y adecuadas, pero si el estudio primario además presenta dimensiones cualitativas tiene ausencias importantes pues convendría que se hicieran notar, aunque no sé si es el caso.</li> <li>– Hay algunos términos de vocabulario que modificaría: La categoría II posiblemente la denominaría "Diseño" más que conformación de grupos. En la categoría VI utilizaría el término "Informe" más que "Reporte" ya que es más común en España. En la categoría V "Recogida de datos" más que "recolección de datos".</li> <li>– Sería aconsejable incluir una categoría sobre el análisis estadístico realizado sobre los datos. Sin embargo, creo que hablaría más de la validez interna y de la de constructo. Los sesgos del investigador no están muy desarrollados ni tampoco las cuestiones de medida. Me parece muy importante reflejar qué datos aportan para poder calcular una magnitud del efecto, sea cual sea su formato (correlación o diferencia de medias).</li> </ul>

- NO. No son adecuadas. Yo utilizaría 3 o a la sumo 4, considerando, como apunto en los márgenes, los criterios generales de valoración de Cook y Campbell.
- En la parte I “Intervention” agregar un indicador: the fidelity of the intervention is provided.
- En la parte IV “Sample”, agrega un indicador: the statistical power of the study (worth working out as could be a major interpretative factor).
- Se echa en falta alguna dimensión sobre el procedimiento de análisis seguido en la resolución del diseño.
- NO, I would include a specific section about the statistical analysis.
- The questions are well chosen and grouped in a logical fashion to facilitate review.
- No. Se debe incluir la categoría diseño de investigación y si el mismo es adecuado para evaluar el efecto de una variable sobre otra (por ejemplo, no sería adecuado un diseño que aún teniendo un grupo experimental y un grupo control tuviese sólo post-test). Es importante preguntar sobre si en la investigación se han controlado las variables extrañas relevantes y si en el estudio se reporta adecuadamente cuáles variables extrañas se controlaron y cómo se controlaron.

### Anexo 15. Proceso de modificación de las dimensiones e ítems de la escala según juicio de expertos

A continuación se explica paso a paso cada uno de los cambios asumidos en las dimensiones e ítems de la escala según las recomendaciones ofrecidas por los expertos.

En la versión inicial del instrumento, en la *dimensión intervención*, el ítem 1.1 “La intervención se sustenta en teoría clara y bien articulada” ha sido valorado positivamente en cuanto al orden (96,4%), contenido (82,1%), pertinencia (83,3%) y redacción (73,3%). Solo se ha realizado un cambio menor en la redacción final con la intención de ser mas precisos en el lenguaje “El diseño de la intervención se sustenta en un marco teórico claro y bien articulado” (ver tabla 1).

Tabla 1 Versión inicial y final ítem 1.1

Ítem 1.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	8	22		1	27		5	23		5	25
	26,7%	73,3%		3,6%	96,4%		17,9%	82,1%		16,7%	83,3%
Versión inicial del ítem				Versión final del ítem				Referencias			
La intervención se sustenta en teoría clara y bien articulada				El diseño de la intervención se sustenta en un marco teórico claro y bien articulado				CONSORT (2010), Morrison, Sullivan, Murray & Jolly (1999), Ramos-Álvarez y Catena (2004), STARD (2003),			
Redacción:											
<ul style="list-style-type: none"><li>– En este ítem puede ser complicado contestar de forma dicotómica Si y No</li><li>– Eliminar la palabra Intervención y sustituirla por Validez de constructo. Esta valoración debería incluir la calidad general de la VI y de la VD. La intervención se apoya en una teoría clara y bien articulada.</li><li>– Entiendo que sería mejor una escala Likert, que permita graduar su respuesta.</li><li>– What clearly means here, each coder can understand that differently</li><li>– These categories need to be defined clearly</li><li>– Se sugiere colocar: datos de identificación, propósito, presencia/ausencia. Al final de la definición de la valoración de la intervención propone colocar "tratamiento realizado". Se sugiere enumerar los ítems</li></ul>											

Fuente: elaboración propia

En el ítem 1.2 “Las variables principales del estudio están claramente definidas” se evidencia un alto porcentaje de acuerdo entre los expertos con respecto a la redacción (86,7%), orden (100%), contenido (82,1%) y pertinencia (93,3%). Se sugiere cualitativamente mejorar el ítem en cuanto a redacción y concreción. Por tanto, los nuevos ítems han atendido ambas solicitudes y se han elaborado con la finalidad de valorar con mayor precisión los estudios primarios en cuanto a la variable dependiente e independiente: “La variable independiente está claramente definida a nivel teórico” y “La variable dependiente está claramente definida a nivel teórico” y “Todos los instrumentos de medición que se han aplicado para medir la variable dependiente se corresponden con su definición conceptual” (ver tabla 2).



Tabla 2 Versión inicial y final ítem 1.2

Ítem 1.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	4 13,3%	26 86,7%			28 100%		5 17,9%	23 82,1%		2 6,7%	28 93,3%
Versión inicial del ítem				Versión final del ítem				Referencias			
Las variables principales del estudio están claramente definidas				La variable independiente está claramente definida a nivel teórico				Chacón, Sanduvete y Alarcón (2005), Shadish, Cook & Campbell (2002)			
				La variable dependiente está claramente definida a nivel teórico				Chacón, Sanduvete y Alarcón (2005), Shadish, Cook & Campbell (2002)			
				Todos los instrumentos de medición que se han aplicado para medir la variable dependiente se corresponden con su definición conceptual				Shadish, Cook & Campbell (2002)			
Redacción											
– Especificar que es "claramente definidas".											
– ¿Variables principales?											
– Not only the description is important, also whether all relevant variables were accounted for.											
– Se sugiere describir la intervención en cuanto a cómo se desarrolla y quién la ha recibido.											
– Yo especificaría: las variables independientes y dependientes.											
Orden:											
– Move down to 4th in this category.											
Contenido											
– I would incorporate some criteria here. E.g, did the authors report particular citations of the theory. In terms of what exactly, demographics, content.											

Fuente: elaboración propia

El porcentaje obtenido en el orden (100%), la pertinencia (89,7%) y el contenido (81,55) del ítem 1.3 “Describe la intervención en cuanto a qué se administra y quién la ha recibido” señalan que es un indicador que ha sido valorado de manera positiva por más del 80% de los expertos. Sin embargo, un poco más de la mitad de los expertos (65,5%) consideran que es necesario hacer mejoras en su redacción. La principal observación es que mide más de un indicador a la vez.

Por ello se ha reformulado en tres ítems independientes (ver tabla 3). El primero tiene que ver con la persona que lleva a cabo la intervención “Ofrece información sobre quién es la persona que lleva a cabo la intervención”. Los otros dos, con el tiempo de duración del programa y su distribución en el tiempo “Ofrece información detallada sobre la intervención en cuanto a número, frecuencia y duración de cada sesión” y “El tiempo de duración de toda la intervención es mayor de 12 semanas (inclusive)”

Tabla 3 Versión inicial y final ítem 1.3

Ítem 1.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,9%	8 27,6%	19 65,5%			27 100%		5 18,5%	22 81,5%		3 10,3%	26 89,7%
Versión inicial del ítem			Versión final del ítem			Referencias					
Describe la intervención en cuanto a qué se administra y quién la ha recibido			Ofrece información sobre quién es la persona que lleva a cabo la intervención			CONSORT (2010), Cook & Campbell (1979), Des Jarlais, Lyles, Crepaz & the TREND group (2004),					
			Ofrece información sobre la intervención en cuanto a número, frecuencia y duración de cada sesión			CONSORT (2010), Cook & Campbell (1979), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Morrison, Sullivan, Murray & Jolly (1999),					
			El tiempo de duración de toda la intervención es mayor de 17 semanas (inclusive)			Slavin (2008)					
Redacción: – Dos preguntas en una. – Revisar la redacción en cuanto a cómo se aplica la variable... – Entiendo que se refiera a "quien" la administra. – La utilización del verbo administrar suena a dar algún tipo de sustancias. – Es redundante poner "describe la intervención en cuánto", mejor decir algo como "se especifica qué se administra y quién ha recibido la intervención". – Se describe suficientemente la intervención (qué se administra, cuánto tiempo, quién la ha recibido). – What is "it?" – Describe cómo la intervención es diferente en el grupo control que en el grupo experimental. – Yo eliminaría "quién la ha recibido" puesto que esto corresponde a la categoría "muestra".											

Fuente: elaboración propia

El ítem 1.4 “Describe cómo la intervención difiere del grupo control” se sugiere ser mejorado en cuanto a su redacción porque casi la mitad de los expertos lo valora como “parcialmente adecuado” (41,4%). Por consiguiente, se ha modificado su redacción a “La variable independiente está claramente definida a nivel operacional porque describe con detalle cómo el grupo de intervención difiere del grupo control” (ver tabla 4)

Tabla 4 Versión inicial y final ítem 1.4

Ítem 1.4											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
1 3,4%	12 41,4%	16 55,2%	1 3,7%	1 3,7%	25 92,6%	1 3,7%	4 14,8%	22 81,5%	1 3,4%	5 17,2%	23 79,3%
Versión inicial del ítem				Versión final del ítem				Referencias			
Describe cómo la intervención difiere del grupo control				La variable independiente está claramente definida a nivel operacional porque describe con detalle cómo el grupo de intervención difiere del grupo control				Chacón, Sanduvete y Alarcón (2005), Shadish, Cook & Campbell (2002)			
Redacción – ¿Intervención difiere del grupo control?											

<ul style="list-style-type: none"> <li>– Se describe</li> <li>– Se describe como la intervención difiere del grupo control</li> <li>– Mejor: Describe cómo el grupo intervención difiere del grupo control</li> <li>– It?</li> <li>– Overly detailed</li> </ul>
Contenido
– In terms of what exactly, demographics, content.

Fuente: elaboración propia

El ítem 1.5 “La intervención se ha llevado a cabo en la universidad o espacios vinculados a la comunidad educativa en donde los propios profesores regulares imparten clases. No se ha llevado a cabo la intervención en aulas especializadas creadas y diseñadas por los investigadores para los propósitos del estudio” ha sido mejorado en cuanto a su redacción porque solo el 33,3% de los expertos lo considera como “totalmente aceptable”. La mayoría de los comentarios cualitativos apuntan a escoger entre la primera y segunda oración del ítem pero no ambas porque crea confusión. La versión final del ítem ha sido “La intervención se ha llevado a cabo en aulas ordinarias y no en laboratorios o aulas especialmente diseñadas para el estudio” (ver tabla 5).

Tabla 5 Versión inicial y final ítem 1.5

Ítem 1.5											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
6	14	10		2	26	4	7	17	2	4	24
20%	46,7%	33,3%		7,1%	92,9%	14,3%	25%	60,7%	6,7%	13,3%	80%
Versión inicial del ítem				Versión final del ítem				Referencias			
La intervención se ha llevado a cabo en la universidad o espacios vinculados a la comunidad educativa en donde los propios profesores regulares imparten clases. No se ha llevado a cabo la intervención en aulas especializadas creadas y diseñadas por los investigadores para los propósitos del estudio				La intervención se ha llevado a cabo en aulas ordinarias y no en laboratorios o aulas especialmente diseñadas para el estudio				CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Morrison, Sullivan, Murray & Jolly (1999).			
Redacción											
– Las dos opciones de respuesta pueden ser válidas para decir lo mismo: "Si", se han llevado a cabo en la universidad. 'No', no se ha llevado a cabo en aulas especializadas.											
– Hay 2 ítems en 1. Mejor la primera formulación que la segunda.											
– Se sugiere esta redacción: la intervención se ha desarrollado en el aula ordinaria y no en aulas creadas y diseñadas para el cumplimiento de los objetivos de la investigación.											
– El “no se ha llevado” no queda claro.											
– La segunda parte "No se ha llevado..." La redactaría como: No se ha llevado a cabo en aulas especializadas creadas y diseñadas por los investigadores para los propósitos.											
– La redacción es algo confusa y puede no quedar claro lo que se está preguntando.											
– Al señalar dos casos que se orientan al mismo punto, debería ponerse, en lugar de un punto, algo así como "o bien..."											
– Son dos ítems distintos, en uno mismo. "la intervención..." "No se ha llevado..."											
– ... donde los profesores imparten clases de forma cotidiana. Eliminar el resto.											
– En este caso sugiero seccionar el indicador, dividiéndole de manera que se pueda responder SI o NO y tener claro en qué espacio se dio la intervención.											

- Esta oración pareciera incluir sólo ambientes educativos y centros formales de enseñanza, ¿cierto?
- La intervención se ha llevado a cabo en la universidad o espacios vinculados a la comunidad educativa en donde los propios profesores regulares imparten clases (no se ha llevado a cabo la intervención en aulas especializadas creadas y diseñadas por los investigadores específicamente para los propósitos del estudio).
- Simplificar redacción.
- Mejor: no se ha llevado a cabo la intervención en ningún otro espacio diseñado o creado o diseñado por los investigadores...
- Este enunciado es demasiado largo.
- Reducir
- Do not understand this – do you mean conducted in labs or regular classrooms?
- Too complex.
- ¿Por qué no se consideraría adecuado que la investigación sea de laboratorio? ¿Se trabajará solamente con investigaciones de campo? Lógicamente en esta área predominan las investigaciones de campo pero no veo por qué si se localiza una de laboratorio no se consideraría adecuada.

Fuente: elaboración propia

Siguiendo con la valoración, menos de la mitad de los expertos, un 46,7% no está totalmente de acuerdo con la redacción del ítem 1.6 “Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural porque el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad”. A esta valoración se suman comentarios cualitativos en donde se menciona que este ítem refiere al mismo contenido pero de una manera mucho más global que el ítem 1.3 “Describe la intervención en cuanto a qué se administra y quién la ha recibido” y el ítem 1.4 “Describe cómo la intervención difiere del grupo control”, razones ambas que han incidido en su eliminación (ver tabla 6).

Tabla 6 Versión inicial y final ítem 1.6

Ítem 1.6											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
4	12	14		1	27	1	8	19		2	28
13,3%	40%	46,7%		3,6%	96,4%	3,6%	28,6%	67,9%		6,7%	93,3%
Versión inicial del ítem			Versión final del ítem						Referencias		
Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural porque el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad			Eliminado						Coalition for Evidence-Based Policy (2010), U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003).		
Redacción											
– Ofrece suficiente información para replicar el estudio. (...) porque el procedimiento es lo suficientemente claro... Diría que esta oración no está bien construida.											
– Sustituir "ejecución" quizá por "aplicación".											
– Esta pregunta y la siguiente creo que dan la misma información. No veo ninguna diferencia.											
– Eliminaría "en un ambiente natural" y "lo que permite la ejecución de la intervención con un alto grado de fiabilidad"											
– Muy larga la redacción, resumir.											
– No queda claro si el ambiente natural es un centro educativo, tal como señala en el ítem anterior, una											

comunidad educativa o ambiente natural es cualquier situación de aprendizaje (ejm un grupo de madres que aprenden pautas de crianza en una comunidad).
– Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural (el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad). Puede ser redundante con lo valorado en el punto 1.3. y 1.4 pero este ítem es más global.
– Simplificar la redacción: el contenido de replicación no necesita explicaciones extraordinarias.
– Este enunciado es demasiado largo.
– Aunque es una práctica habitual en los meta-análisis indicar que los estudios seleccionados deben ofrecer SUFICIENTE información para ser replicados podría ser incluido en este estudio una definición más operativizable (más concreta).
– A bit lengthy.
– In similar situations' and 'in a natural environment' might be contradictory: maybe the experiment was in a lab?
– Rephrase: specialized training in not necessary to replicate the intervention.
– Yo eliminaría "lo que permite la ejecución de la intervención con un alto grado de fiabilidad".
– Ofrece suficiente información para replicar el estudio en situaciones similares.
– "Sufficient information is provided to replicate the study".

Fuente: elaboración propia

Este ítem 1.7 “La intervención puede ser replicada por investigadores quienes tienen distintos niveles de formación” presenta las mismas dificultades que el anterior. Más de la mitad de los expertos lo valora en su redacción entre 0 (20%) y 1 (33,3%) siendo la puntuación máxima 2. Las valoraciones cualitativas indican que es repetitivo porque otros como el 1.3 “Describe la intervención en cuanto a qué se administra y quién la ha recibido”, el 1.4 “Describe cómo la intervención difiere del grupo control” y el 1.6 “Ofrece suficiente información para replicar el estudio en situaciones similares en un ambiente natural porque el procedimiento es lo suficientemente claro, lo que permite la ejecución de la intervención con un alto grado de fiabilidad”, ya lo incluyen. Además los expertos consideran que en caso de dejarlo es necesario “operativizarlo”. Por todas estas razones se ha decidido eliminar este ítem por considerarse que se mide a través de otros como los ya mencionados (ver tabla 7).

Tabla 7 Versión inicial y final ítem 1.7

Ítem 1.7											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
6 20%	10 33,3%	14 46,7%	1 3,6%	2 7,1%	25 89,3%	5 17,9%	3 10,7%	20 71,4%	4 13,3%	4 13,3%	22 73,3%
Versión inicial del ítem				Versión final del ítem				Referencias consultadas			
La intervención puede ser replicada por investigadores quienes tienen distintos niveles de formación				Eliminado							
Redacción											
– En español no se suele utilizar el 'quienes'. Se entiende, pero suena raro.											
– No se si es imprescindible.											
– "la intervención puede ser replicada por otros investigadores independientemente de su nivel de formación"											

- Se sugiere eliminar el ítem 1.7 porque: puede ser redundante con lo valorado en el punto 1.3 y 1.4. no es necesario. El 1.6 sería suficiente.
  - Se sugiere introducir en cada dimensión un ítem de valoración global, que puede servir para la toma de decisiones en situaciones ambiguas o para determinar la congruencia del propio evaluador. Por ejemplo: valore globalmente la definición de la variable independiente en una escala de 1 a 10. Por ejemplo: valore globalmente la definición de la variable dependiente en una escala de 1 a 10.
  - Ya se ha hablado de la replicación en la pregunta anterior. Si lo relevante aquí es lo referido al nivel de formación de los investigadores, céntrese la pregunta en ellos.
  - La redacción no es correcta. ¿Qué tal "con investigadores con distintos niveles de formación"?
  - "quienes tienen". Es un ítem poco claro y no bien formulado.
  - Yes, "(...) levels seems same as above one"
  - With different levels of experience or skills.
  - El valor de una investigación no se fundamenta en que pueda ser replicada por investigadores con distintos niveles de formación sino que pueda ser replicada. Opino que con el ítem anterior es suficiente.
  - La intervención puede ser replicada por investigadores con distintos niveles de formación
- Orden
- ¿Cuáles serían los distintos niveles de formación? Pareciera que deberían estar claro los tópicos y las estrategias metodológicas empleadas y verificar si el facilitador las maneja.

Fuente: elaboración propia

Siguiendo, con la estructura de la versión inicial de la escala, en la *dimensión conformación de grupos*, el ítem 2.1 “Contempla al menos un grupo experimental y un grupo control” se ha eliminado. Aunque si bien es cierto el 80% valora su redacción y el 85,7% su contenido con la puntuación mas alta, es imprescindible mencionar que se trata de un criterio de inclusión y no de un ítem que mida la calidad metodológica de los estudios primarios (ver tabla 8). El hecho de que algunos expertos se hayan percatado de este error ha exigido revisar nuevamente los criterios de inclusión con detenimiento.

Tabla 8 Versión inicial y final del ítem 2.1

Ítem 2.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2	4	24		1	27	2	2	24		1	28
6,7%	13,3%	80%		3,6%	96,4%	7,1%	7,1%	85,7%		3,4%	96,6%
Versión inicial del ítem				Versión final del ítem					Referencias		
Contempla al menos un grupo experimental y un grupo control				Eliminado							
Redacción											
– Parece más un ítem de inclusión/exclusión del meta-análisis.											
– Es imprescindible.											
– Yo cambiaría contempla por existe.											
– Antes de valorar los ítems, eliminar "Conformación de grupos" y apuntar: Validez interna y externa del diseño utilizado: valoración de la aleatorización y las técnicas de control empleadas. Luego, directamente sobre el ítem apunta: se contempla al menos un grupo experimental y un grupo control.											
– ¿no forma parte de los criterios de inclusión?											
– What about if anything is not reported or clearly defined? What about if it meets one but not the other. Or if there is contamination that is explained and reported with some numeric information.											
– This would mean that a repeated measures study (with the same group repeated before and after/during a treatment) is getting a low score. It also means that single-case experimental designs would get a low score, although also these designs can be internally valid. Maybe it is therefore better to talk a control and treatment condition, rather than about a control and treatment group.											
– Contempla una división de la "n" en 2 grupos (experimental y control)											
Contenido											

– What if pre and post?

Fuente: elaboración propia

También, el ítem 2.2 ha sido eliminado “Hace referencia a si el grupo control no está conformado por sujetos que tenían la opción de participar en la intervención pero se han negado” porque tanto los comentarios cualitativos como el valor otorgado por los expertos en cuanto a su redacción (valores 0 y 1) apuntan a que más de la mitad, un 51,7% considera que no es adecuado o parcialmente adecuado (ver tabla 9).

Tabla 9 Versión inicial y final del ítem 2.2

Ítem 2.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
6	9	14	1	3	23	5	7	15	1	8	20
20,7%	31%	48,3%	3,7%	11,1%	85,2%	18,5%	25,9%	55,6%	3,4%	27,6%	69%
Versión inicial del ítem				Versión final del ítem				Referencias			
Hace referencia a si el grupo control no está conformado por sujetos que tenían la opción de participar en la intervención pero se han negado				Eliminado							
Redacción											
– No aplicable si no hay grupo control. Tener este ítem en negativo lo hace bastante confuso. Cuando dices 'Hace referencia' te refieres a que este ítem hace referencia a si el grupo control (...) o a que en el artículo se hace referencia a (...). En este segundo caso sería un ítem de reporting.											
– Eso cómo lo sabemos, ¿en todos los estudios primarios se indica?											
– Sustituir el término "conformado" por "compuesto", Además hacer más clara la redacción del ítem para que quede claro lo que se pregunta.											
– La redacción en negativo dificulta la comprensión. ¿En positivo que no esté formado por ese grupo de personas?											
– Es muy confuso, se propone: "El grupo control no está formado por participantes que se han negado a participar en el grupo experimental"											
– El grupo control no está conformado por sujetos que tenían la opción de participar en la intervención pero se han negado. Poner en positivo, se hace poco entendible											
– La utilización de negativos dificulta la comprensión del ítem. Se sugiere: el grupo de control se puede considerar equivalente al grupo experimental antes de la aplicación del tratamiento.											
– La redacción negativa puede dificultar la respuesta. No queda claro a qué se responde con sí o no.											
– Hace referencia a si en el grupo control se descartaron/eliminaron sujetos que tenían la opción a participar en la intervención pero se negaron											
– Formulación oscura: cabría formular el ítem en términos de aleatoriedad, etc. Este ítem se solapa con lo que se formula en III: ASIGNACIÓN											
– confusing											
– I would move this issue to where you talk about random assignment											
– Avoid using the negative, it is confusing											
– NO se entiende lo que se quiere decir con el ítem											
Pertinencia											
– This would be difficult to know											

Fuente: elaboración propia

Asimismo, el ítem 2.3 se ha eliminado “El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención” porque varios expertos

mencionan que será difícil su medición debido a que no resulta ser una condición que se menciona con regularidad en las investigaciones (ver tabla 10).

Tabla 10 Versión inicial y final del ítem 2.3

Ítem 2.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
3	3	22		2	24	1	6	19		4	23
10,7%	10,7%	78,6%		7,7%	92,3%	3,8%	23,1%	73,1%		14,8%	85,2%
Versión inicial del ítem			Versión final del ítem					Referencias			
El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención			Eliminado								
Redacción											
– ¿A qué medidas te refieres? ¿Medidas de respuesta? Si es así, parece que el ítem no acaba de cuadrar en un bloque titulado 'conformación de grupos'.											
– ¿Eso se hace así?											
– No veo claro que vaya en esta categoría de conformación de grupos.											
– El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención.											
– No sé a qué se refiere.											
– How would you know this?											
– It is no clear to me what you mean.											
– El investigador escoge las medidas para el análisis de los resultados antes de administrar la intervención.											
– Esto no es posible saberlo con la información que aparece en un artículo.											
Contenido:											
– I do not see this information is easy to figure out by research papers.											

Fuente: elaboración propia

El ítem 2.4 “Se demuestra que no hay diferencias sistemáticas entre el grupo de intervención y de control antes de la intervención (situación basal). Ambos grupos poseen un alto grado de homogeneidad antes de la intervención en cuanto a: niveles de rendimiento académico, datos demográficos...” se considera importante pero su redacción no ha sido muy bien evaluada porque más de la mitad de los expertos lo valora entre 0 (26,7%) y 1 (26,7%). El comentario más repetitivo se centra en especificar si la homogeneidad refiere al rendimiento académico y si la equivalencia viene dada estadísticamente. La redacción final ha sido “Antes de la intervención, se demuestra estadísticamente que el grupo control y experimental son inicialmente equivalentes en cuanto a la variable dependiente” (ver tabla 11).

Tabla 11 Versión inicial y final del ítem 2.4

Ítem 2.4											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
8	8	14	1	3	24	3	4	21	1	1	28
26,7%	26,7%	46,7%	3,6%	10,7%	85,7%	10,7%	14,3%	75%	3,3%	3,3%	93,3%



Versión inicial del ítem	Versión final del ítem	Referencias
Se demuestra que no hay diferencias sistemáticas entre el grupo de intervención y de control antes de la intervención (situación basal). Ambos grupos poseen un alto grado de homogeneidad antes de la intervención en cuanto a: niveles de rendimiento académico, datos demográficos...	Antes de la intervención, se demuestra estadísticamente que el grupo control y experimental son inicialmente equivalentes en cuanto a la variable dependiente	CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Shadish, Cook & Campbell (2002).
<p>Redacción</p> <ul style="list-style-type: none"> <li>– No aplicable si no hay grupo control 'un alto grado de homogeneidad' es una expresión bastante ambigua. ¿Tal vez se podría operativizar diciendo que no hay diferencias estadísticamente significativas? Los puntos suspensivos dejan el ítem demasiado abierto. ¿Qué ocurre si valoran esos dos niveles pero en uno hay diferencias y en otro no? ¿Qué ocurre si en estos dos hay homogeneidad pero no en un tercer nivel? ¿Qué ocurre si no se han valorado estos niveles?</li> <li>– ¿Eso es tan importante? No se entienden las dos oraciones, y no es adecuado poner puntos suspensivos...</li> <li>– Creo que sería más correcto decir "significativas" que "sistemáticas". De nuevo, en lugar de punto pondría "o bien..."</li> <li>– No hay diferencias semánticas entre el grupo de intervención y de control antes de la intervención, ya que ambos grupos poseen un alto grado de homogeneidad en cuanto a: niveles de rendimiento académico, datos demográficos...</li> <li>– Lo dividiría en dos indicadores.</li> <li>– Algunos de estos aspectos demográficos pueden no estar nombrados, entonces se pondría que no porque se asumen que al nombrarlo, es que no se consideran. ¿Cierto?</li> <li>–Este ítem sobraría si se reformula el ítem 2.2.</li> <li>– Simplificar la redacción.</li> <li>– Este enunciado es demasiado largo.</li> <li>– Incluir "o se controlan en los análisis de datos las diferencias existentes".</li> <li>– Simplify.</li> <li>– Possible differences can also be controlled for in the statistical analysis.</li> <li>– Proven is a strong statement. There was no obvious bias in composition of the control and experimental groups.</li> <li>– Se demuestra que los grupos experimental y control son homogéneos antes de la intervención.</li> <li>– Lo clave es que los grupos de control y experimental sean equivalentes en cuanto al rendimiento académico (la variable dependiente). Esto implica que en el diseño debe tener un pretest. Si los grupos difieren significativamente en el pretest es fundamental que se analice la magnitud de la diferencia. Otros aspectos como el sexo, la edad, etc son variables que deben ser controladas, pero esto debería estar como ítem aparte.</li> </ul> <p>Contenido</p> <ul style="list-style-type: none"> <li>– Both groups are highly homogeneous before the intervention with regard to: academic performance levels, demographic data.</li> <li>– I would include here some criteria about the way the authors have tested that homogeneity.</li> </ul>		

Fuente: elaboración propia

El ítem 2.5 “El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo” ha sido valorado en su redacción con la máxima puntuación por el 80% de los expertos y con un porcentaje muy parecido, con un 78,6% en su contenido. La otra observación que se ha atendido es que son dos indicadores en un mismo ítem por lo cual se ha dividido en dos indicadores en la versión final “Todos los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y

experimental” y “Todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales” (ver tabla 12)

Tabla 12 Versión inicial y final del ítem 2.5

Ítem 2.5											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
3	3	24			28	3	3	22		2	28
10%	10%	80%			100%	10,7%	10,7%	78,6%		6,7%	93,3%
Versión inicial del ítem				Versión final del ítem				Referencias			
El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo				Todos los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y experimental				Campbell & Stanley (1963), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Shadish, Cook & Campbell (2002).			
				Todos los instrumentos de medida que se han aplicado en el grupo control y experimental son iguales							
Redacción											
– No aplicable si no hay grupo control. ¿Qué responder si se ha recogido de la misma manera pero al mismo tiempo (o viceversa)?											
– Eso es muy importante											
– Se recogen los resultados del grupo de intervención y control de la misma manera y al mismo tiempo											
– "de la misma manera" es poco preciso											
– What about if it meets one but not the other?											
– The dependent variable is measured in a similar way (same instrument, same conditions) in intervention and control groups											
– Yo lo redactaría así: en el estudio se recogen los datos del grupo experimental y control de la misma manera y al mismo tiempo.											

Fuente: elaboración propia

Las puntuaciones menos altas del ítem 2.6 “Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental, es decir, no hay evidencias de contaminación (*cross-over*)” han sido en la redacción con un 70% y en el contenido con un 71,4%. Por tanto, se ha mejorado la redacción final del ítem “Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental (contaminación *cross-over*)” (ver tabla 13).

Tabla 13 Versión inicial y final del ítem 2.6

Ítem 2.6											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
4	5	21	1	1	26	6	2	20	1	3	26
13,3%	16,7%	70%	3,6%	3,6%	92,9%	21,4%	7,1%	71,4%	3,3%	10%	86,7%
Versión inicial del ítem			Versión final del ítem			Referencias					
Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental, es decir, no hay evidencias de contaminación ( <i>cross-over</i> )			Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental (contaminación <i>cross-over</i> )			Cook & Campbell (1979), Shadish, Cook & Campbell (2002)					

Redacción
– No aplicable si no hay grupo control. Dos preguntas en una: Podría ser que ningún sujeto del grupo control ha participado posteriormente en el grupo experimental' pero aún así podría haber evidencias de cross-over. También podría haber algún sujeto del grupo control que haya participado pero que no se vean afectados (y no se considere cross-over). Finalmente, el cross-over podría producirse también porque un participante del grupo experimental finalmente haya participado en el grupo control. ¿Podrían los participantes del grupo control conocer la intervención recibida por el grupo expuesto y beneficiarse de ella sin que haya cross-over?
– Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental.
– Una cuestión menor: ¿grupo control o grupo de control? Ambas son correctas, pero sería deseable que se emplease la misma acepción a lo largo de la escala.
– Evitar la formulación con términos negativos ("ningún").
– "Ningún sujeto del grupo de control".
– But wedge designs do this and they are excellent.
– Ok, but this rarely is a problem.
– "No participant of the control group has later in the experimental group".
Orden
– Muy importante.
Contenido
– Or if there is contamination that is explained and reported with some numeric information.

Fuente: elaboración propia

Pasando ahora a la *dimensión asignación*, el ítem 3.1 “Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental” ha sido valorado con la máxima puntuación en la redacción por el 75,9% de los expertos, en el orden con un 92,6%, en el contenido con un 85,2% y en la pertinencia con un 86,2%. No obstante, para mejorar la redacción y así evitar que se confunda este ítem con el siguiente, el 3.2 “Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental”, se ha cambiado la palabra “aulas (secciones)” por “grupos” quedando redactado de esta manera “Se han seleccionado aleatoriamente los grupos que van a participar en el estudio” (ver tabla 14).

Tabla 14 Versión inicial y final del ítem 3.1

Ítem 3.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,9%	5 17,2%	22 75,9%		2 7,4%	25 92,6%	1 3,7%	3 11,1%	23 85,2%		4 13,8%	25 86,2%
Versión inicial del ítem				Versión final del ítem				Referencias			
Las aulas (secciones) han sido asignadas aleatoriamente como grupo control o experimental				Se han seleccionado aleatoriamente los grupos que van a participar en el estudio				CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos-Álvarez y Catena (2004), Sánchez-Meca, López-Pina y Nuñez-Nuñez (2012).			
Redacción											
– No aplicable sino hay grupo control.											
– Quitar secciones.											
– Estos dos ítems se confunden, pareciera ser el mismo.											
– Para la parte III. Asignación: se sigue valorando la calidad del estudio en cuando a su valor. Puede haber diferencias pero dudo que, así redactado el ítem, el evaluador las pueda captar. Se puede dejar éste pero entonces quitaría el anterior.											

<ul style="list-style-type: none"> <li>– Modificar la definición de asignación "(...)" que afecten a los resultados".</li> <li>– Cambiar la palabra assignation por assignment.</li> <li>– This and the next seem redundant. Rephrase.</li> <li>– Yo eliminaría este ítem</li> </ul> <p>Contenido</p> <ul style="list-style-type: none"> <li>– Why classrooms? If individuals, rather than groups or classrooms are assigned to the conditions, the study would get a low score on this topic although assignment of individuals is preferable from a statistical point of view.</li> </ul>
---

Fuente: elaboración propia

Al haber cambiado la redacción del ítem anterior, la redacción del ítem 3.2 “Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental” no sufrió modificaciones. El hecho de que ambos ítems se relacionen entre sí ha sido motivo principal para que la redacción no haya sido valorada tan positivamente con un 76,7%, no así el orden (92,9%), el contenido (85,7%) y la pertinencia (86,2%) (ver tabla 15).

Tabla 15 Versión inicial y final del ítem 3.2

Ítem 3.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
4	3	23	1	1	26	2	2	24	2	2	25
13,3%	10%	76,7%	3,6%	3,6%	92,9%	7,1%	7,1%	85,7%	6,9%	6,9%	86,2%
Versión inicial del ítem			Versión final del ítem			Referencias					
Los grupos han sido asignados aleatoriamente como grupo control o experimental			Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental			CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos-Álvarez y Catena (2004), Sánchez-Meca, López-Pina y Nuñez-Nuñez (2012).					
Redacción											
– ¿Diferencia con el ítem anterior?											
– ¿Qué diferencia hay entre aula y grupo?											
– Uniría los 3 ítems poniendo. Las unidades (aulas, grupos o sujetos) han sido asignados aleatoriamente al grupo control o experimental.											
– Estos dos ítems se confunden, pareciera ser lo mismo.											
– "Me parece igual al ítem 2.6 de la segunda parte".											
– "Parece utilizarse el término "grupo" en 2 sentidos (como grupo de sujetos y como grupo experimental o de control). En todo caso, aunque se entiende qué se quiere decir, la formulación no es clara.											
– Same as above.											
– Yo lo redactaría así: "los tratamientos han sido asignados aleatoriamente a los grupos". Yo colocaría este ítem como el segundo.											
Contenido											
– What is the difference with the previous topic? If you say yes on the previous topic, you also have to say yes here...											

Fuente: elaboración propia

Como se ha comentado en párrafos anteriores, se espera que exista una clara diferenciación entre los ítems 3.1 y 3.2, diferenciación que también se exige al ítem 3.3 “Los sujetos han sido asignados aleatoriamente al grupo control o experimental” por estar todos estrechamente relacionados entre sí. Sin embargo, de los tres ítems, el 3.3 es

el que mejor ha sido evaluado en cuanto a su redacción y pertinencia, ambos criterios con un 90% de aprobación (ver tabla 16).

Tabla 16 Versión inicial y final del ítem 3.3

Ítem 3.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	3 10%	27 90%	1 3,6%		27 96,4%		1 3,6%	27 96,4%	1 3,3%	2 6,7%	27 90%
Versión inicial del ítem				Versión final del ítem				Referencias			
Los sujetos han sido asignados aleatoriamente al grupo control o experimental				Los sujetos han sido asignados aleatoriamente al grupo control y experimental				CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos–Álvarez y Catena (2004), Sánchez–Meca, López–Pina y Nuñez–Nuñez (2012).			
Redacción											
– Estas tres cuestiones están íntimamente relacionadas y creo que inducen a confusión.											
– Redundant with the questions of category II.											
– Yo colocaría este ítem de primero. Lo redactaría así: "los sujetos han sido asignados aleatoriamente a los grupos". Si hay asignación aleatoria de los sujetos a los grupos, entonces no se estaría trabajando con aulas (secciones) tal y como están conformadas naturalmente. Si se responde que no a este ítem entonces ya se sabe que los grupos son naturales y aquí la clave es que los tratamientos sean asignados aleatoriamente a los grupos.											

Fuente: elaboración propia

El ítem 3.4 “No se ha usado la aleatorización para la asignación de sujetos a los grupos pero se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación” no ha sido muy bien valorado en su redacción por el 60% de los expertos pero sí en orden (96,4%), contenido (78,6%) y pertinencia (93,3%%). En consecuencia, se ha modificado a “Se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación”, descripción que aunque se presente en la tabla 17 no corresponde a la valoración máxima del ítem con 1 punto sino a un valor de 0,5 (ver anexo 7 que corresponde a la versión final de la escala, el ítem 6.3 en donde se muestran los distintos descriptores para cada valoración).

Tabla 17 Versión inicial y final del ítem 3.4

Ítem 3.4											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,7%	10 33,3%	18 60%		1 3,6%	27 96,4%	3 10,7%	3 10,7%	22 78,6%		2 6,7%	28 93,3%
Versión inicial del ítem				Versión final del ítem				Referencias			

No se ha usado la aleatorización para la asignación de sujetos a los grupos pero se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación)	Se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación*	CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Sánchez-Meca, López-Pina y Nuñez-Nuñez (2012).
<p>Redacción</p> <ul style="list-style-type: none"> <li>– 'Si': se ha aplicado algún procedimiento de control. 'No?': Se ha aplicado aleatorización. (Falta una opción de no aplicable) Basta con que haya una variable controlada (p. ej. Género) o se tienen que haber controlado un mínimo de variables como para considerar que no hay sesgo? No hay paréntesis de apertura'.</li> <li>– No entiendo bien el alcance.</li> <li>– "Aunque no se ha utilizado... se ha aplicado algún otro procedimiento..."</li> <li>– Al final debe poner "punto". No cierre de paréntesis.</li> <li>– Quitar paréntesis final</li> <li>– Se repite en el ítem anterior.</li> <li>– Cambio el orden para que no empiece con un no. Queda así: Se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación, si no se ha podido usar la aleatorización para la asignación de sujetos a los grupos.</li> <li>– Este ítem forma parte del anterior. Si la escala es politómica entonces no obtendría el máximo valor</li> <li>– Tal y como está formada esta categoría, parece que ningún estudio podrá obtener la puntuación máxima de calidad. Esto no supone un gran problema, aunque sería bueno que los futuros usuarios fuesen conscientes de este hecho.</li> <li>– No to the above is the same</li> <li>– Strange?</li> <li>– Not sure about the intended meaning of this. What's wrong with randomization? It seems to contradict the following item.</li> <li>– NO se ha usado la aleatorización para la asignación de sujetos a los grupos, pero se ha aplicado algún procedimiento de control de las variables extrañas, como por ejemplo: emparejamiento, bloqueo o estratificación</li> </ul> <p>Contenido</p> <ul style="list-style-type: none"> <li>– Where would you include quasi-experimental designs.</li> </ul>		

\* este ítem no corresponden en la versión final a la puntuación máxima, es decir, al valor 1

Fuente: elaboración propia

Por último, el ítem 3.5 “Advierte sobre cualquier factor que pudiera comprometer el proceso de aleatorización” ha sido eliminado porque las valoraciones cualitativas de los expertos apuntan a que es reiterativo y que se sobreentiende que la aleatorización ya implica el control de factores (ver tabla 18).

Tabla 18 Versión inicial y final del ítem 3.5

Ítem 3.5											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2	6	21		1	26	2	6	19	2	6	21
6,9%	20,7%	72,4%		3,7%	96,3%	7,4%	22,2%	70,4%	6,9%	20,7%	72,4%
Versión inicial del ítem			Versión final del ítem				Referencias				
Advierte sobre cualquier factor que pudiera comprometer el proceso de aleatorización			Eliminado								
Redacción											
– En otras escalas se pregunta explícitamente por el proceso de aleatorización. No sé si es habitual en tu campo indicarlo...											
– Eso es muy importante, fundamental.											
– "Los autores informan sobre cualquier otro factor que pudiera comprometer el proceso de											

aleatorización".
– No se entiende lo que se quiere decir, especificar.
– Advierte sobre cualquier factor que pudiera comprometer el proceso de aleatorización. En caso afirmativo descríbalos...
– Este ítem se podría eliminar si sólo se considera que existe asignación aleatoria cuando no hay factores que puedan alterarla.
– Examples here.
– Maybe give some examples, or be more clear?
– No se entiende qué se quiere decir con este ítem.

Fuente: elaboración propia

Continuando con el juicio de expertos, en la versión original del instrumento, en la *dimensión muestra*, el ítem 4.1 “Describe con detalles los criterios de inclusión de la muestra” el 86,7% de los expertos considera que la redacción es totalmente adecuada así como el orden (96,4%), la pertinencia (93,3%) y el contenido (85,7%) obtuvieron valoraciones positivas. Los dos únicos cambios obedecen a la redacción y al contenido quedando así su redacción final “Menciona las características de la muestra” (ver tabla 19).

Tabla 19 Versión inicial y final del ítem 4.1

Ítem 4.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
1 3,3%	3 10%	26 86,7%		1 3,6%	27 96,4%		4 14,3%	24 85,7%		2 6,7%	28 93,3%
Versión inicial del ítem			Versión final del ítem					Referencias			
Describe con detalles los criterios de inclusión de la muestra			Menciona las características de la muestra					CONSORT (2010), Morrison, Sullivan, Murray & Jolly (1999), STARD (2003).			
Redacción: – Faltaría colocar la definición de población. – Se describen... – Describe con detalles los criterios de selección de la muestra. – Los criterios de inclusión se aplican a cada sujeto que compone la muestra. – "(...) que afecten a los resultados" – Reformular "Describe con detalles los criterios de inclusión Y DE EXCLUSIÓN de la muestra" – Examples or details about that "detail" – There should also be information on the way the sample is drawn: by preference a random sample. Maybe also on whether the representativeness of the sample for the population was checked.											

Fuente: elaboración propia

El ítem 4.2 “Incluye el número de grupos que han sido escogidos del total de la población para luego especificar el tamaño de la muestra” se ha redefinido en el ítem 3.1 en donde se valora el procedimiento de la selección de la población “Se han seleccionado aleatoriamente los grupos que van a participar en el estudio”. De esta forma se ha atendido al 29,6% restante (valores 0 y 1) de los expertos quienes consideraron que habría que mejorar la redacción (ver tabla 20).

Tabla 20 Versión inicial y final del ítem 4.2

Ítem 4.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
3 11,1%	5 18,5%	19 70,4%		1 4%	24 96%	1 4%	2 8%	22 88%	1 3,7%	4 14,8%	22 81,5%
Versión inicial del ítem			Versión final del ítem					Referencias			
Incluye el número de grupos que han sido escogidos del total de la población para luego especificar el tamaño de la muestra			Se han seleccionado aleatoriamente los grupos que van a participar en el estudio					CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), STARD (2003).			
<div>– Redacción poco clara. Por el ítem ("Los sujetos han sido asignados aleatoriamente al grupo control o experimental") parece que la unidad de medida también podría ser el sujeto, no?</div> <div>– Creo que este ítem hay que especificar más para tener claro lo que se está preguntando</div> <div>– No se entiende</div> <div>– Se incluye...</div> <div>– Muy extenso, o se resume o se subdivide</div> <div>– It is not clear to me what you mean</div> <div>– Incluye el número de grupos que han sido escogidos del total de la población, para luego especificar el tamaño de la muestra</div> <div>– Yo más bien preguntaría sobre si se describe el procedimiento de selección de la muestra y si dicho procedimiento es adecuado. Sería importante incluir un ítem sobre la adecuación del tamaño de la muestra.</div>											

Fuente: elaboración propia

El ítem 4.3 refiere a mortalidad diferencial entre el grupo control y experimental "El estudio incluye un análisis que asegura que el desgaste de la muestra (si lo hay) no socava la equivalencia de los grupos de intervención y de control. El porcentaje de participantes que se ha perdido durante el proceso de recolección de datos es lo suficientemente pequeño que no afecta los resultados del grupo de intervención y de control (no supera el 80% de la muestra original)" ha sido reformulado con especial atención debido a que mas de la mitad de los expertos, un 60%, lo valora entre 0 y 1. Se ha reducido su extensión y se han atendido los comentarios relativos al contenido. La redacción final del ítem ha sido "El porcentaje de participantes que ha iniciado el estudio se mantiene al terminar la investigación, no hay pérdida de sujetos" y "El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación" (ver tabla 21).

Tabla 21 Versión inicial y final del ítem 4.3

Ítem 4.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
5 16,7%	13 43,3%	12 40%	1 3,6%	1 3,6%	26 92,9%	1 3,6%	5 17,9%	22 78,6%		2 6,7%	28 93,3%
Versión inicial del ítem			Versión final del ítem			Referencias					



El estudio incluye un análisis que asegura que el desgaste de la muestra (si lo hay) no socava la equivalencia de los grupos de intervención y de control. El porcentaje de participantes que se ha perdido durante el proceso de recolección de datos es lo suficientemente pequeño que no afecta los resultados del grupo de intervención y de control (no supera el 80% de la muestra original)	El porcentaje de participantes que ha iniciado el estudio se mantiene al terminar la investigación, no hay pérdida de sujetos	Chacón, Sanduvete y Alarcón (2005), CONSORT (2010), Shadish, Cook & Campbell (2002), STARD (2003).
	El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación	
<p>Redacción</p> <ul style="list-style-type: none"><li>– Dos preguntas en una.</li><li>– Intentar expresar de forma más clara lo que se pregunta.</li><li>– Al igual que en los casos anteriores, se hacen dos afirmaciones que no han de yuxtaponerse sino reforzarse. "O bien...", "Es decir", "En otras palabras"</li><li>– Es demasiado largo, la primera frase no aporta más información que la segunda.</li><li>– No se entiende bien. Poner de forma clara y menos extensa.</li><li>– Simplificar la redacción.</li><li>– "No supera el 80% de la muestra original" resulta confuso. Describir más bien entre paréntesis cuál sería "el porcentaje suficientemente más pequeño que no afecta los resultados del grupo experimental y del grupo control.</li><li>– Este enunciado es demasiado largo. Además, un 80% de mortalidad es claramente una tasa demasiado alta. ¿Quizá se refiere a un 20% de mortalidad, con lo cual tenemos un 80% de los participantes después de la intervención?</li><li>– 2 ítem en 1</li><li>– "sample wear (not were)"</li><li>– ¿attribution??</li><li>– What do you mean with the sample wear?</li><li>– Awkwardly phrased. Try to make this shorter.</li><li>– "(...) recolección de los datos es lo suficientemente pequeño que no afecta a los (...)"</li></ul> <p>Contenido</p> <ul style="list-style-type: none"><li>– Me parece preferible la expresión "pérdida muestral" que "desgaste de la muestra"</li></ul>		

Fuente: elaboración propia

El ítem 4.4 “Se ha solicitado a los sujetos de la muestra su consentimiento para participar en la investigación antes de conocer si estaban asignados al grupo experimental o control” se ha eliminado porque es no es nada frecuente que esto se mencione en las investigaciones. Se ha tomado esta decisión por los comentarios cualitativos ofrecidos por los expertos aunque también es cierto que ha sido valorado por ellos como adecuado por más del 79,3% en cuanto a su redacción, 80% de pertinencia, 96,3% en orden y 92,6% en el contenido (ver tabla 22).

Tabla 22 Versión inicial y final del ítem 4.4

Ítem 4.4											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
1 3,4%	5 17,2%	23 79,3%		1 3,7%	26 96,3%	1 3,7%	1 3,7%	25 92,6%	3 10%	3 10%	24 80%
Versión inicial del ítem			Versión final del ítem			Referencias					

Se ha solicitado a los sujetos de la muestra su consentimiento para participar en la investigación antes de conocer si estaban asignados al grupo experimental o control	Eliminado	Coalition for Evidence-Based Policy (2010), Shadish, Cook & Campbell (2002), U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003).
<p>Redacción</p> <ul style="list-style-type: none"> <li>– No se diferencia entre no haber solicitado su consentimiento y haberlo hecho después de la asignación. Si bien afecta a la ética de la investigación, ¿en qué medida afecta a la calidad de los resultados?</li> <li>– Quizás sería interesante preguntar si los sujetos participan voluntariamente en el estudio</li> <li>– "(...) grupo experimental o de control"</li> <li>– Why is this important?</li> <li>– Por lo que se dice en un artículo no es posible que el lector sepa si el consentimiento informado fue solicitado antes de la asignación del sujeto a uno de los grupos. Además, este ítem hace referencia al cumplimiento de las normas éticas de investigación, no a la categoría "muestra".</li> </ul> <p>Contenido</p> <ul style="list-style-type: none"> <li>– The statistical power of the study (worth working out as could be a major interpretative factor)</li> </ul>		

Fuente: elaboración propia

Siguiendo, con la estructura de la versión inicial y cómo sean realizado los cambios en función de la valoración de los expertos, en la *dimensión recolección de datos*, el ítem 5.1 “Los instrumentos utilizados miden con exactitud los resultados reales por los cuales la intervención está destinada a incidir. Existe una correspondencia entre los instrumentos y la intervención realizada” ha obtenido una valoración muy baja por parte de los expertos. Menos de la mitad, un 48,3% considera que este ítem cumple con una adecuada redacción (ver tabla 23).

A esto se suma algunos comentarios cualitativos que refieren a que este indicador ya había sido mencionado el ítem 1.2 “Las variables principales del estudio están claramente definidas” y en el ítem 2.5 “El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo”. Estas dos razones han sido decisivas para la eliminación de este indicador sin dejar de recordar que su contenido está presente en la versión final de la escala en los indicadores “La variable independiente está claramente definida a nivel teórico”, “La variable dependiente está claramente definida a nivel teórico”, “Todos los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y experimental” y “Todos los instrumentos de medida que se han aplicado en el grupo control y experimental son iguales (respectivamente, ver anexo 7 los indicadores 2.1, 2.3, 9.1 y 9.2 que corresponden a la versión final de la escala).

Tabla 23 Versión inicial y final del ítem 5.1

Ítem 5.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
3 10,3%	12 41,4%	14 48,3%		2 7,4%	25 92,6%	1 3,7%	5 18,5%	21 77,8%		4 13,8%	25 86,2%
Versión inicial del ítem				Versión final del ítem				Referencias			
Los instrumentos utilizados miden con exactitud los resultados reales por los cuales la intervención está destinada a incidir. Existe una correspondencia entre los instrumentos y la intervención realizada				Eliminado				Chacón, Sanduvete y Alarcón (2005), Shadish, Cook & Campbell (2002)			
Redacción											
– Redacción poco clara. Poco claro. 'Exactitud' me hace pensar en fiabilidad. 'reales' me hace pensar en validez. Valorar si existe una correspondencia entre los instrumentos y la intervención me parece que es otra cosa, no?											
– Redactar mejor.											
– Pondría... "resultados reales sobre los cuales la intervención..."											
– Revisar la redacción: los instrumentos utilizados miden con exactitud la variable sobre la que se espera que incida la intervención...											
– Igual que el anterior.											
– Ya ha sido mencionado en la categoría “intervención”.											
– Sobra la primera frase, con las segunda se entiende el objetivo											
– ¿Qué significa miden con exactitud? ¿fiables? ¿válidos? Especificar. Yo quitaría: existe una correspondencia entre los instrumentos y la intervención realizadas											
– Estás preguntando dos cosas en este ítem: 1) si el instrumento tiene validez de constructo o contenido sobre lo que se pretende medir. Y 2) si lo que dice que mide el instrumento coincide con lo que el investigador quiere incidir con la intervención.											
– Con respecto al tópico de recolección de datos: yo lo englobaría con el apartado I bajo la rúbrica "Validez de constructo" porque aquí se trata fundamentalmente de como se define operativamente la VD (lo mismo que en el apartado I se hace con la VI).											
– Simplificar la redacción.											
– Se repite en el punto 1.2 y 2.5 cuando dice si existe relación entre la VD y su definición conceptual y operacional.											
– ¿No sería mejor: "... los resultados reales en los cuales la intervención..."?											
– La respuesta está en dos partes: 1) "The instruments measure precisely the dependent variable on which...". 2) It is not clear to me what you mean with that correspondence											
– "(...) Existe una correspondencia entre los instrumentos de evaluación o medición para la intervención realizada"											
Contenido											
– Include something about blinded and if the people that take the measure are the same one or not to those who do the intervention.											

Fuente: elaboración propia

El ítem 5.2 “El grupo control y experimental han sido evaluados con los mismos instrumentos” ha sido valorado positivamente por los expertos en pertinencia (92,9%), orden (92,3%), redacción (85,7%) y contenido (84,6%) al considerar que están “totalmente de acuerdo” con el indicador (ver tabla 24). Sin embargo, el comentario asertivo de uno de los expertos indica que este ítem se repite porque ya ha sido mencionado en el ítem 2.5 “El estudio recoge los resultados del grupo de intervención y

control de la misma manera y al mismo tiempo”. Por tanto, como se ha comentado en párrafos anteriores, se ha redactado nuevamente mejorando su redacción “Todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales” (ver tabla 24).

Tabla 24 Versión inicial y final del ítem 5.2

Ítem 5.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	4	24		2	24	2	2	22	1	1	26
	14,3%	85,7%		7,7%	92,3%	7,7%	7,7%	84,6%	3,6%	3,6%	92,9%
Versión inicial del ítem			Versión final del ítem					Referencias			
El grupo control y experimental han sido evaluados con los mismos instrumentos			Todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales					Campbell & Stanley (1963), Chacón, Sanduvete y Alarcón (2005), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Shadish, Cook & Campbell (2002).			
Redacción											
– En un ítem anterior ya se pregunta: "El estudio recoge los resultados del grupo de intervención y control de la misma manera y al mismo tiempo"											
– ¿Y si hay diferentes momentos de medida?											
– Está muy clara la pregunta, pero si no cumple con estos entonces no es un grupo control											
– Al igual que en el ítem 2.5, se repite											
– Creo que ésta es una cuestión obvia y prescindible (al menos, a juzgar por los estudios que se publican en mi campo)											
– 2 ítems en 1. Los grupos de control y experimental.											
– So what I wrote under it is supposed to be here											
– This could be stated more succinctly											
– Los grupos "control y experimental (...)"											

Fuente: elaboración propia

El ítem 5.3 “Todos los instrumentos aplicados para medir rendimiento han sido validados/estandarizados” ha sido modificado por “Todos los instrumentos de medición que se han aplicado para medir la variable dependiente han sido validados/estandarizados”. Más del 70% de los expertos han coincidido en valorar como “totalmente de acuerdo” este ítem en cuanto al orden (96,3%), pertinencia (89,7%), redacción (75,9%) y contenido (74,1%) (ver tabla 25).

Tabla 25 Versión inicial y final del ítem 5.3

Ítem 5.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
1	6	22		1	26	2	5	20		3	26
3,4%	20,7%	75,9%		3,7%	96,3%	7,4%	18,5%	74,1%		10,3%	89,7%
Versión inicial del ítem			Versión final del ítem			Referencias					
Todos los instrumentos aplicados para medir rendimiento han sido validados/estandarizados			Todos los instrumentos de medición que se han aplicado para medir la variable dependiente han sido			Des Jarlais, Lyles, Crepaz & the TREND group (2004), Shadish, Cook & Campbell (2002).					

	validados/estandarizados	
<p>Redacción</p> <ul style="list-style-type: none"> <li>– Quizá se podría incluir "se aportan datos sobre la fiabilidad y validez de los instrumentos utilizados"</li> <li>– En la palabra rendimiento se sugiere: "La variable dependiente"</li> <li>– Es una cuestión importante, aunque quizá aquí estamos siendo demasiado exigentes. Es posible que existan estudios que combinan medidas estandarizadas (e.g. un test) con otras no estandarizadas (e.g., un diario de los participantes). En mi opinión, esto no debería ser penalizado en la escala.</li> <li>– Sugiero evitar "validados/estandarizados"</li> <li>– An excellent example where a Yes or No will not suffice.</li> <li>– They may use some that did and some that did not, so then? Should it say that there should be at least 2 like that. How?</li> <li>– What is meant by 'validated'? What if a validation study was performed, but the validity was not very good? What about the reliability of the instrument?</li> </ul>		

Fuente: elaboración propia

El ítem 5.4 resulta muy extenso “Las medidas de resultado no están a favor del grupo de intervención sobre el grupo control o viceversa. Por ejemplo, un estudio de un programa computarizado para enseñar matemáticas a los jóvenes estudiantes no deben medir los resultados mediante un examen computarizado, ya que el grupo de intervención probablemente tendrá mayor facilidad con el ordenador que el grupo control”. Por ello solo el 64,3% de los expertos considera que está totalmente de acuerdo con la redacción aunque si hay una valoración más positiva en cuanto a su contenido (92,3%), orden (88,5%) y pertinencia (85,7%). La versión final del ítem luego de la revisión de las valoraciones cualitativas ha sido “Los instrumentos de medida no dan ventaja al grupo experimental” (ver tabla 26).

Tabla 26 Versión inicial y final del ítem 5.4

Ítem 5.4											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
3	7	18	2	1	23	2		24	1	3	24
10,7%	25%	64,3%	7,7%	3,8%	88,5%	7,7%		92,3%	3,6%	10,7%	85,7%
Versión inicial del ítem				Versión final del ítem				Referencias			
Las medidas de resultado no están a favor del grupo de intervención sobre el grupo control o viceversa. Por ejemplo, un estudio de un programa computarizado para enseñar matemáticas a los jóvenes estudiantes no deben medir los resultados mediante un examen computarizado, ya que el grupo de intervención probablemente tendrá mayor facilidad con el ordenador que el grupo control				Los instrumentos de medida no dan ventaja al grupo experimental				Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Shadish, Cook & Campbell (2002).			
Redacción											
– Mejorar quizás la redacción											
– Las medidas... no favorecen de antemano...											

- "Se informa sobre las características psicométricas de los instrumentos utilizados"
- Demasiado texto...
- No entiendo la pregunta
- Yo lo redactaría de otra manera, porque luego hay que poner un ejemplo muy largo.
- Simplificar la redacción
- Puede ser interesante tener en cuenta este aspecto, pero el enunciado es indudablemente demasiado largo.
- ¿? Ítem demasiado largo; si se decide mantenerlo inclúyase un paréntesis (Por ejemplo...)
- Are not "biased" in favour
- I added 'biased' because it is possible that the intervention group will score better, but exactly due to the intervention affect you are interested in.

Fuente: elaboración propia

En la versión inicial, la *dimensión resultados* en donde se valora la interpretación y discusión de los hallazgos del estudio, el ítem 6.1 "Describe cómo la intervención ha afectado los resultados" se sugiere integrar con el ítem 6.2 "Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos" (ver tablas 27 y 28, respectivamente) porque son muy similares en su contenido. Por tanto, la versión final de este ítem se ha centrado en la intervención y no tanto en el informe de resultados "Se discuten las causas por las cuales la intervención ha sido efectiva o no" (ver tabla 27).

Tabla 27 Versión inicial y final del ítem 6.1

Ítem 6.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2	5	23		1	27		4	24	1	2	27
6,7%	16,7%	76,7%		3,6%	96,4%		14,3%	85,7%	3,3%	6,7%	90%
Versión inicial del ítem			Versión final del ítem					Referencias			
Describe cómo la intervención ha afectado los resultados			Se discuten las causas por las cuales la intervención ha sido efectiva o no					CONSORT (2010), Ramos–Álvarez y Catena (2004), STARD (2003).			
Redacción											
– No me puedo imaginar un estudio de intervención en que no se comuniquen los efectos de la intervención.											
– Se describe...											
– Donde está el tópico IV Reporte de resultados, sugiere titularlo: Validez de la conclusión estadística. Hay que incluir algún ítem que haga referencia a la adecuación de los análisis realizados ¿se han comprobado supuestos del modelo estadístico? ¿se ajustan los análisis realizados a los objetivos de la investigación y a la naturaleza de los datos? Son cuestiones relativas a la validez de la conclusión estadística que pueden afectar enormemente a la calidad de la investigación.											
– Es muy parecido al siguiente ítem.											
– Buscar unificar este ítem y el siguiente porque son similares.											
– It seems very ambiguous.											
– This is vague. What do you mean with 'how'? To what degree, what underlying processes...?											
– En mi opinión, aquí es clave que en la investigación se realicen todas las comparaciones que deben realizarse en los diseños pretest–posttest, y no sólo los contrastes postes–posttest, como sucede en muchos casos. Lo primero es que indique si la intervención tuvo efecto y luego cuál fue la dirección de ese efecto. Yo lo redactaría así: "describe cómo la intervención ha afectado al rendimiento"											

Fuente: elaboración propia

Y como se ha comentado en el párrafo anterior, el ítem 6.2 “Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos” se ha fusionado con el ítem precedente para así no duplicar información porque los comentarios cualitativos apuntan a esta misma idea (ver tabla 28).

Tabla 28 Versión inicial y final del ítem 6.2

Ítem 6.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
1	4	24			27	1	2	24	4	1	25
3,4%	13,8%	82,8%			100%	3,7%	7,4%	88,9%	13,3%	3,3%	83,3%
Versión inicial del ítem			Versión final del ítem					Referencias			
Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos			Se discuten las causas por las cuales la intervención ha sido efectiva o no					CONSORT (2010), Ramos–Álvarez y Catena (2004).			
Redacción											
– Está bien pero que algo sea fácil y sencillo es subjetivo de acuerdo a los conocimientos previos que se tengan. Esto sólo para estar pendiente de ítems como éste.											
– Yo no incluiría este ítem. ¿Qué entendemos por "claros y sencillos"? La valoración de este ítem puede diferir drásticamente en función de la base metodológica y experiencia investigadora del evaluador.											
– Se repite con el anterior.											
– Es igual que el anterior. Unificar en un solo ítem.											
– Se parece al ítem anterior.											
– You can list some minimum stats that should be reported.											
– "Informa de los efectos de la intervención de una manera fácil para su comprensión, utilizando (...)"											
Orden											
– La comunicación del estudio (reporting) no tiene por qué estar relacionada con la calidad del estudio											
– Se...											

Fuente: elaboración propia

El ítem 6.3 “Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos que solo han obtenido un efecto positivo” ha sido valorado con la menor calificación en cuanto a su pertinencia, un 92,9%, dato que exigió su eliminación de la escala (ver tabla 29).

Tabla 29 Versión inicial y final del ítem 6.3

Ítem 6.3											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
		8	22		1	27		2	26	1	3
		26,7%	73,3%		3,6%	96,4%		7,1%	92,9%	3,3%	10%
Versión inicial del ítem			Versión final del ítem			Referencias					

Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos que solo han obtenido un efecto positivo	Eliminado	CONSORT (2010), Morrison, Sullivan, Murray & Jolly (1999), Ramos-Álvarez y Catena (2004).
<p>Redacción</p> <ul style="list-style-type: none"> <li>– Informa de los resultados tanto favorables como desfavorables.</li> <li>– Se repiten contenidos, son dos ítems muy parecidos.</li> <li>– ¿Mejor:... respecto a todos los resultados que el estudio ha medido y no sólo respecto a los positivos.</li> <li>– Sugiere que la redacción de la pregunta sea: It... with regard to all the outcome variables the study... an not only on positive effects.</li> <li>– Igual a los dos anteriores. Los contenidos son similares.</li> <li>– Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos de los que sólo han obtenido un efecto positivo.</li> <li>– Modificar la redacción de"... a todos los resultados que el estudio ha medido".</li> </ul> <p>Orden</p> <ul style="list-style-type: none"> <li>– No sólo de aquellos que han obtenido un efecto positivo quitar ese "solo".</li> </ul>		

Fuente: elaboración propia

En cuanto al ítem 6.4 “Incluye los resultados principales a los cuales responde la pregunta de investigación así como también resultados secundarios de interés” se ha modificado su redacción y su contenido aunque ha obtenido por parte de los expertos, un 75,9% y 92,6% de aprobación respectivamente.

Esto obedece a que los comentarios cualitativos de los ítems 6.1 “Describe cómo la intervención ha afectado los resultados”, 6.2 “Informa de los efectos de la intervención de una manera fácil para su comprensión utilizando términos claros y sencillos que permiten valorar la importancia de los resultados obtenidos” y 6.3 “Informa de los efectos de la intervención con respecto a todos los resultados que el estudio ha medido y no sólo de aquellos que solo han obtenido un efecto positivo” coinciden en que los tres indicadores son muy parecidos entre sí y que además no necesariamente miden la calidad de los estudios porque están orientados hacia los resultados. Al revisar nuevamente la literatura, el cambio se ha focalizado en la implementación de la intervención. La versión final intenta atender estas deficiencias “Informa de las dificultades para la implementación de la intervención” (ver tabla 30).

Tabla 30 Versión inicial y final del ítem 6.4

Ítem 6.4											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	7	22			27		2	25	4	1	24
	24,1%	75,9%			100%		7,4%	92,6%	13,8%	3,4%	82,8%
Versión inicial del ítem			Versión final del ítem			Referencias					



Incluye los resultados principales a los cuales responde la pregunta de investigación así como también resultados secundarios de interés	Informa de las dificultades para la implementación de la intervención	CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Ramos–Álvarez y Catena (2004).
Redacción – No tengo claro en cómo esto puede afectar a la calidad de un estudio. – Informa tanto de los resultados principales relacionados con el planteamiento del problema como de... – Esto me parecen dos ítems en realidad. – "Informa sobre los datos"... Debe mejorarse la redacción del ítem. – Sugiere que la pregunta sea: ...results if there are of interest. – Incluye los resultados principales a los que se responde la pregunta de investigación así como las resultados secundarios de interés.		

Fuente: elaboración propia

El último ítem de esta dimensión en la versión inicial del ítem 6.5 “Reporta los datos de incluso aquellos miembros del grupo de intervención que no participaron o completaron la intervención (*Intention to treat analysis*) como de aquellos sujetos que terminaron la intervención (*Completers analysis*)” que ha obtenido una valoración positiva en cuanto a su contenido (92, 3%), orden (96,2%), pertinencia (78,6%) y redacción (71,4%). Sin embargo, se ha modificado mejorando su contenido y redacción “La estrategia de análisis ha sido por intención de tratar (*intention to treat analysis*)” (ver tabla 31).

Tabla 31 Versión inicial y final del ítem 6.5

Ítem 6.5											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
3	5	20		1	25		2	24	2	4	22
10,7%	17,9%	71,4%		3,8%	96,2%		7,7%	92,3%	7,1%	14,3%	78,6%
Versión inicial del ítem			Versión final del ítem					Referencias			
Reporta los datos de incluso aquellos miembros del grupo de intervención que no participaron o completaron la intervención ( <i>Intention to treat analysis</i> ) como de aquellos sujetos que terminaron la intervención ( <i>Completers analysis</i> )			La estrategia de análisis ha sido por intención de tratar ( <i>intention to treat analysis</i> )					Chacón, Sanduvete y Alarcón (2005), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Sánchez–Meca, López–Pina y Nuñez–Nuñez (2012), Shadish, Cook & Campbell (2002).			
Redacción											
– Dos preguntas en una.											
– Mejorar la redacción, no sé si es fundamental.											
– Informa en lugar de reporta.											
– Tiene relación con la pregunta anterior. Pudiera utilizarse un solo ítem.											
– Esto para mí es nuevo, yo sólo estaba atenta a que reportaran los que terminaron, pero si es un criterio, me imagino que no todos lo cumplirán.											
– Sugiero introducir en cada apartado un ítem de valoración global, que puede servir para la toma de decisiones en situaciones ambiguas o para determinar la congruencia del propio evaluador.											
– Revisar la redacción.											
– Este enunciado puede resultar confuso. El <i>intention–to–treat analysis</i> y el <i>completers analysis</i> son en realidad dos alternativas diferentes. Yo eliminaría la segunda del texto, de manera que el indicador de calidad sea haber realizado un análisis por intención de tratar.											
– Añadiría un ítem como una fortaleza del estudio si se han hecho seguimientos de al menos 3 meses de los resultados añadiría otro ítem valorando como fortaleza del estudio si los datos de los resultados se											

pueden convertir en tamaños del efecto.  
 – Incluso reporta los datos de aquellos miembros del grupo de intervención que no participaron o complementaron la intervención (*intention to treat analysis*) así como de aquellos sujetos que terminaron la intervención (*completers analysis*). No solo reporta los datos de los grupos que terminaron la intervención sino también de aquellos que no participaron o no completaron la intervención  
 – ¿Cómo se pueden incluir los datos de quienes no participaron?

Fuente: elaboración propia

Para cerrar con las modificaciones de los ítems, la *dimensión sesgo del investigador*, estuvo conformada por dos ítems. En el ítem 7.1 “El mismo investigador es quien lleva a cabo la intervención en el grupo experimental y control” se ha modificado ligeramente su redacción “Es la misma persona quien lleva a cabo la intervención tanto en el grupo control como experimental”. La redacción (76,7%), el contenido (78,6%), la pertinencia (86,7%) y el orden (92,9%) han sido valorados en su mayoría con la máxima puntuación (ver tabla 32).

Tabla 32 Versión inicial y final del ítem 7.1

Ítem 7.1											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
	7	23		2	26		6	22	1	3	26
	23,3%	76,7%		7,1%	92,9%		21,4%	78,6%	3,3%	10%	86,7%
Versión inicial del ítem			Versión final del ítem					Referencias			
El mismo investigador es quien lleva a cabo la intervención en el grupo experimental y control			Es la misma persona quien lleva a cabo la intervención tanto en el grupo control como experimental					Chacón, Sanduvete y Alarcón (2005), Shadish, Cook & Campbell (2002).			
Redacción											
– Suele ocurrir pero podría no ofrecer información.											
– No queda claro si se refiere a que es la misma persona quien realiza la intervención en ambos grupos o si se refiere al investigador en comparación con otra persona como puede ser el profesor del aula.											
– Sugiero introducir en cada apartado un ítem de valoración global, que puede servir para la toma de decisiones en situaciones ambiguas o para determinar la congruencia del propio evaluador.											
– Añadir "o se controlan las posibles diferencias entre las distintas personas que aplican las intervenciones (e.g. se hacen manuales las intervenciones).											
– Possibly there is no 'intervention' in the group. Moreover, maybe there is not really a roe of the researcher in the intervention.											
– El mismo investigador es quien lleva a cabo la intervención.											

Fuente: elaboración propia

Y el ítem 7.2 “Quien aplica la intervención desconoce los objetivos del estudio y en consecuencia a qué grupo pertenece el sujeto que está evaluando (control ciego)” ha sido dividido en dos para atender individualmente el enmascaramiento de los sujetos y de los evaluadores (ver tabla 33).

Tabla 33 Versión inicial y final del ítem 7.2

Ítem 7.2											
Redacción			Orden			Contenido			Pertinencia		
0	1	2	0	1	2	0	1	2	0	1	2
2 6,9%	3 10,3%	24 82,8%	1 3,7%		26 96,3%	1 3,7%	3 11,1%	23 85,2%	1 3,4%	2 6,9%	26 89,7%
Versión inicial del ítem				Versión final del ítem				Referencias			
Quien aplica la intervención desconoce los objetivos del estudio y en consecuencia a qué grupo pertenece el sujeto que está evaluando (control ciego)				Los sujetos desconocen si forman parte del grupo control o experimental				Chacón, Sanduvete y Alarcón (2005), CONSORT (2010), Sánchez-Meca, López-Pina y Nuñez-Nuñez (2012), Shadish, Cook & Campbell (2002), STARD (2003).			
				Los evaluadores desconocen los objetivos de la investigación							
Redacción											
– Una cosa no implica necesariamente la otra.											
– Estaría bien que fuera así.											
– Y en consecuencia... si el sujeto está asignado al grupo experimental o al de control											
– Enmascaramiento, la palabra "ciego" está en desuso por sus connotaciones. Quitar el paréntesis.											
– Si el que aplica la intervención es el propio investigador sí debe conocer los objetivos, lo que podría no saber, sobre todo si la medición es individual a qué grupo pertenece. Esto a veces no es reportado en el paper de la investigación, igual que pregunte antes, si no se reporta se asume que es no o se pudiera agregar una tercera opción de respuesta que fuera (NR o no se reporta).											
– Es interesante incluir este aspecto, aunque aquí se están mezclando conceptos: desconocer los objetivos de la investigación y desconocer la pertenencia grupal de cada sujeto del estudio son situaciones distintas, y no tienen por qué darse simultáneamente.											
– And what about the 'blinding' participant?											
– Entiendo por lo que se dice respecto a cómo se puntuará si el evaluador responde "Si" o "No" a cada uno de los ítems que lo adecuado es que sea el mismo investigador quien administre la intervención lo cual efectivamente es así. Por ende, es imposible que quien aplica la intervención desconozca los objetivos del estudio y a qué grupo pertenece el sujeto que está evaluando.											

Fuente: elaboración propia

Ahora bien, elaborados todos los cambios y eliminaciones de los ítems según las valoraciones de los expertos, también es cierto que estas propiciaron la creación de otros indicadores.

Un grupo nuevo de ítems se centró en aspectos vinculados a la *validez de constructo* relativos a las consideraciones teóricas, objetivos e hipótesis del estudio, entre ellos: “Se mencionan estudios previos en donde la intervención es eficaz a través de diseños aleatorios los cuales se han implementado en contextos similares” y “Los objetivos e hipótesis del estudio son claros”. Otros atendieron asuntos relacionados con la difusión/imitación del tratamiento o con la rivalidad compensatoria de los sujetos: “Se garantiza que los sujetos del grupo control y experimental no comparten información o intentan copiar el tratamiento” y “Se compensa al grupo control con técnicas cualitativas (como entrevistas no estructuradas y observación directa) para evitar que busquen competir con el grupo experimental”. Y, un último indicador focalizado en el la medición de la VD por un solo método “La variable dependiente es medida con mas de un método” (ver tabla 34).

Tabla 34 Nuevos ítems relacionados con la validez de constructo

Nuevos ítems	Referencias
Se mencionan estudios previos en donde la intervención es eficaz a través de diseños aleatorios los cuales se han implementado en contextos similares	CONSORT (2010), Ramos-Álvarez y Catena (2004), STARD (2003).
Los objetivos e hipótesis del estudio son claros	CONSORT (2010), Des Jarlais, Lyles, Crepaz & the TREND group (2004), Education group on guidelines on evaluation (1999), Morrison, Sullivan, Murray & Jolly (1999), Ramos-Álvarez y Catena (2004).
Se garantiza que los sujetos del grupo control y experimental no comparten información o intentan copiar el tratamiento	Cook & Campbell (1979), Cook, Campbell & Peracchio (1990)
Se compensa al grupo control con técnicas cualitativas (como entrevistas no estructuradas y observación directa) para evitar que busquen competir con el grupo experimental	Campbell (2002), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Ramos-Álvarez y Catena (2004), Shadish, Cook &
La variable dependiente es medida con mas de un método	Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Shadish, Cook & Campbell (2002),

Fuente: elaboración propia

Por otra parte, en cuanto a la *validez interna*, la recomendación ofrecida por uno de los expertos de incluir no solo un indicador que hablara acerca de las características de la muestra sino otro relacionado con los criterios de inclusión y exclusión ha servido para la creación un nuevo ítem “Explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos”. Asimismo, los comentarios realizados en cuanto a la mortalidad global de la muestra han dado pie para la elaboración de este otro ítem “El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación” (ver tabla 35).

Tabla 35 Nuevos ítems relacionados con la validez interna

Nuevos ítems	Referencias
Explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos	Shadish, Cook & Campbell (2002), STARD (2003).
El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación	Chacón, Sanduvete y Alarcón (2005), CONSORT (2010), Shadish, Cook & Campbell (2002), Ramos-Álvarez y Catena (2004).

Fuente: elaboración propia

Y, a este listado se suman otros dos. El primero, “Un año después de finalizada la intervención, se mide nuevamente la variable dependiente” por considerar que se vincula con otros que refieren al momento de medición de la VD y la homogeneidad de los instrumentos de medida. El segundo y último, concerniente a los incentivos “Los

sujetos participan en la investigación de manera voluntaria sin recibir recompensas” (ver tabla 36).

Tabla 36 Nuevos ítems relacionados con la validez interna

Nuevos ítems	Referencias
Un año después de finalizada la intervención, se mide nuevamente la variable dependiente	Chacón, Sanduvete y Alarcón (2005), Des Jarlais, Lyles, Crepaz & the TREND group (2004).
Los sujetos participan en la investigación de manera voluntaria sin recibir recompensas	Shadish, Cook & Campbell (2002).

Fuente: elaboración propia

Igualmente se crearon ítems para atender la *validez externa* en función de los comentarios de los expertos y debido a la revisión de otras referencias bibliográficas. La interpretación y generalización de resultados ha sido el norte (ver tabla 37).

Tabla 37 Nuevos ítems relacionados con la validez externa

Nuevos ítems	Referencias
Los resultados de la intervención se discuten en función de la teoría y estudios empíricos relevantes	Ramos-Álvarez y Catena (2004), STARD (2003).
A partir de los hallazgos encontrados se sugieren posibles investigaciones futuras	Ramos-Álvarez y Catena (2004).
Se generalizan los resultados de la intervención solo a la población que ha sido estudiada	Campbell & Stanley (1963), CONSORT (2010), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Ramos-Álvarez y Catena (2004).
Se generalizan los resultados de la intervención a situaciones o contextos que han sido estudiados	CONSORT (2010), Cook & Campbell (1979), Cook, Campbell & Peracchio (1990), Ramos-Álvarez y Catena (2004).

Fuente: elaboración propia

Por último, ya para finalizar este apartado, en la tabla 38 se presentan los ítems vinculados a la *validez de conclusión estadística*, los cuales han sido sugeridos por varios expertos en las distintas valoraciones cualitativas. Esta dimensión recoge aspectos vinculados con el tamaño de la muestra “El tamaño de la muestra tanto en el grupo control como experimental es mayor de 30 sujetos”, la fiabilidad de los instrumentos de medición “La fiabilidad de todos los instrumentos de medición es mayor a 0,80”, los supuestos del modelo estadístico “Las pruebas estadísticas aplicadas tienen en cuenta los supuestos que deben cumplir los datos”, la magnitud del efecto “Ofrece información con respecto a la magnitud del efecto” y el nivel de confianza “El nivel de confianza es de 99%”

Tabla 38 Nuevos ítems validez de conclusión estadística

Nuevos ítems	Referencias
El tamaño de la muestra tanto en el grupo control como experimental es mayor de 30 sujetos	U.S. Department of Education Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance (2003)

La fiabilidad de todos los instrumentos de medición es mayor a 0,80	Nunnally & Bernstein (1994)
Las pruebas estadísticas aplicadas tienen en cuenta los supuestos que deben cumplir los datos	Ramos-Álvarez, Moreno-Fernández, Valdés-Conroy & Catena (2008), STARD (2003).
Ofrece información con respecto a la magnitud del efecto	CONSORT (2010), Promising practices network (2013), STARD (2003).
El nivel de confianza es de 99%	Promising practices network (2013)

Fuente: elaboración propia

**Anexo 16. Versión final de la escala de calidad metodológica según recomendaciones del juicio de expertos**

**Código:**

**Fecha:**

**Evaluador:**

Información general documento primario	
Nombre los autores	
Título del documento	
Año	

Criterios de inclusión	Si	No
VI: programa aprendizaje cooperativo y aprendizaje individual (grupo control inactivo)		
VD: rendimiento ( <i>outcome</i> )		
Estudiantes universitarios de todos los niveles		
Al menos un grupo control y experimental		
Con diseño pre-test y pos-test o solo post-test		
Suficiente información para el cálculo de la magnitud del efecto (ME)		

**PUNTUACIÓN POR TIPO DE VALIDEZ**

Tipos de validez	Valor total	%	Total
Validez de constructo	MAX 13	30%	
Validez interna	MAX 19	30%	
Validez externa	MAX 7	10%	
Validez de conclusión estadística	MAX 5	30%	
<b>Total</b>	<b>44</b>	<b>100%</b>	

**CALIDAD METODOLÓGICA**

Niveles de calidad	Valor
Bajo (entre 0 y 4, 99 ambos inclusive)	
Medio (entre 5 y 6,99 ambos inclusive)	
Alto (entre 7 y 10, ambos inclusive)	

## VALIDEZ DE CONSTRUCTO

1. CONSIDERACIONES TEÓRICAS, OBJETIVOS E HIPÓTESIS DEL ESTUDIO					
Indicador	0		0,5	1	Valor
1.1 Teoría	El diseño de la intervención se sustenta en un marco teórico débil que refleja un bajo dominio en el área de estudio dejando afuera una gran cantidad de aspectos relevantes de la teoría		El diseño de la intervención se sustenta en un marco teórico que refleja conocimiento del área de estudio pero deja de mencionar algunos aspectos relevantes de la teoría	El diseño de la intervención se sustenta en un marco teórico claro y bien articulado	
1.2 Estudios previos	Se mencionan estudios previos en donde la intervención es eficaz a través de diseños no aleatorios los cuales se han implementado en contextos no similares		Se mencionan estudios previos en donde se demuestra que la intervención es eficaz a través de diseños aleatorios y/o no aleatorios en contextos similares	Se mencionan estudios previos en donde la intervención es eficaz a través de diseños aleatorios los cuales se han implementado en contextos similares	
	0			1	
1.3 Objetivos e hipótesis	Los objetivos e hipótesis del estudio no son claros			Los objetivos e hipótesis del estudio son claros	
2. DEFINICIÓN DE CONSTRUCTOS					
Indicador	0		1		
2.1 Definición teórica VI	La variable independiente no está claramente definida a nivel teórico			La variable independiente está claramente definida a nivel teórico	
2.2 Definición operacional VI	La variable independiente no está claramente definida a nivel operacional porque describe solo el grupo de intervención sin explicar cómo difiere del grupo control			La variable independiente está claramente definida a nivel operacional porque describe con detalle cómo el grupo de intervención difiere del grupo control	
2.3 Definición teórica VD	La variable dependiente no está claramente definida a nivel teórico			La variable dependiente está claramente definida a nivel teórico	
	0		0,5	1	
2.4 Adecuación instrumentos de medida VD	Más de un instrumento de medida o el único que se ha aplicado para medir la variable dependiente no se corresponde con su definición conceptual		Al menos uno de todos los instrumentos que han sido aplicados para medir la variable dependiente se corresponde con su definición conceptual	Todos los instrumentos de medición que se han aplicado para medir la variable dependiente se corresponden con su definición conceptual	
	0		0,5	1	
2.5 Validación instrumentos de medida VD	No hace mención a si los instrumentos que se han aplicado para medir la variable dependiente han sido validados/estandarizados	Más de un instrumento de medida o el único que se ha aplicado para medir la variable dependiente no ha sido validado/estandarizados	Al menos uno de todos los instrumentos que han sido aplicados para medir la variable dependiente no ha sido validado/estandarizados	Todos los instrumentos de medición que se han aplicado para medir la variable dependiente han sido validados/estandarizados	



## VALIDEZ DE CONSTRUCTO (continuación)

3. SESGOS OPERACIONALIZACIÓN CAUSA-EFECTO/REACTIVIDAD SUJETOS				
Indicador	0		1	
<b>3.1</b> Sesgo único método medición VD	La variable dependiente es medida solo a través de un mismo método		La variable dependiente es medida con más de un método	
<b>3.2</b> Interacción administración de pruebas y tratamiento	Los instrumentos de medida dan ventaja al grupo experimental		Los instrumentos de medida no dan ventaja al grupo experimental	
	0		1	
<b>3.3</b> Difusión/imitación del tratamiento	No ofrece información con respecto a si los participantes del grupo control y experimental comparten o no información o intentan copiar el tratamiento	Los sujetos del grupo control y experimental comparten información o intentan copiar el tratamiento	Se garantiza que los sujetos del grupo control y experimental no comparten información o intentan copiar el tratamiento	
<b>3.4</b> Interacción de tratamientos intra-sujetos	No ofrece información que demuestre que ninguno de los sujetos del grupo control ha participado posteriormente en el grupo experimental	Sujetos del grupo control han participado posteriormente en el grupo experimental (contaminación –cross-over–)	Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental (contaminación cross-over)	
<b>3.5</b> Rivalidad compensatoria de los sujetos	No ofrece información sobre las técnicas que se utilizan para evitar que el grupo control intente competir con el grupo experimental	No se compensa al grupo control con técnicas cualitativas para evitar que busquen competir con el grupo experimental	Se compensa al grupo control con técnicas cualitativas (como entrevistas no estructuradas y observación directa) para evitar que busquen competir con el grupo experimental	
TOTAL VALIDEZ DE CONSTRUCTO				

## VALIDEZ INTERNA

4. DISEÑO INVESTIGACIÓN					
Indicador	0		0,5	1	
4.1 Situación basal (homogeneidad)	No ofrece información con respecto a si el grupo control y experimental son inicialmente equivalentes en cuanto a la variable dependiente	Menciona que el grupo control y experimental no son inicialmente equivalentes en cuanto a la variable dependiente	Antes de la intervención, se describen las características demográficas de los participantes en el grupo control y experimental como evidencia de que ambos son inicialmente equivalentes	Antes de la intervención, se demuestra estadísticamente que el grupo control y experimental son inicialmente equivalentes en cuanto a la variable dependiente	
5. CRITERIOS INCLUSIÓN/EXCLUSIÓN MUESTRA					
Indicador	0		1		
5.1 Características muestra	No menciona las características de la muestra		Menciona las características de la muestra		
5.2 Inclusión–exclusión de sujetos	0		1		
	No explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos		Explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos		
6. PROCEDIMIENTO SELECCIÓN/ASIGNACIÓN MUESTRA					
Indicador	0		1		
6.1 Selección grupos (población)	No ofrece información sobre cómo han sido seleccionados los grupos (población)	No se han seleccionado aleatoriamente los grupos que van a participar en el estudio	Se han seleccionado aleatoriamente los grupos que van a participar en el estudio		
6.2 Asignación grupos	No ofrece información sobre cómo han sido asignados el grupo control y el grupo experimental	Los grupos no han sido asignados aleatoriamente como grupo control o grupo experimental	Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental		
	0		0,5	1	
6.3 Asignación sujetos	No ofrece información sobre cómo han sido asignados los sujetos al grupo control y al grupo experimental	Los sujetos no han sido asignados aleatoriamente al grupo control y experimental	Se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación	Los sujetos han sido asignados aleatoriamente al grupo control y experimental	

## VALIDEZ INTERNA (continuación)

7. MORTALIDAD							
Indicador	0		0,5		1		
7.1 Mortalidad (sujetos del estudio)	El porcentaje de participantes que no ha llegado a terminar el estudio supera el 20% de la muestra original		El porcentaje de participantes que no ha llegado a terminar el estudio es entre 0 y 20% de la muestra original (ambos inclusive)		El porcentaje de participantes que ha iniciado el estudio se mantiene al terminar la investigación, no hay pérdida de sujetos		
	0		0,25	0,5	1		
7.2 Mortalidad (grupo control y experimental)	El porcentaje de la diferencia de pérdidas de participantes entre el grupo control y experimental es mayor del 20%		El porcentaje de la diferencia de pérdida de participantes entre el grupo control y experimental es entre el 0 y 20%	El porcentaje de la diferencia de pérdida de participantes entre el grupo control y experimental es igual	El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación		
8. ADMINISTRACIÓN DEL TRATAMIENTO							
Indicador	0			1			Valor
8.1 Persona que administra el tratamiento	No ofrece información sobre quién es la persona que lleva a cabo la intervención			Ofrece información sobre quién es la persona que lleva a cabo la intervención			
	0		0,5		1		
8.2 Lugar del tratamiento	No ofrece información sobre el lugar en donde se ha llevado a cabo la intervención		La intervención se ha llevado a cabo en laboratorios o aulas especialmente diseñadas para el estudio		La intervención se ha llevado a cabo en aulas ordinarias y no en laboratorios o aulas especialmente diseñadas para el estudio		
8.3 Distribución duración tratamiento	No ofrece información del número, frecuencia y duración de cada sesión	Ofrece solo información del número, frecuencia o duración de cada sesión	Ofrece información en dos de los aspectos relacionados con la duración de la intervención: número, frecuencia o duración de cada sesión		Ofrece información sobre la intervención en cuanto a número, frecuencia y duración de cada sesión		
	0	0,25	0,5	0,75	1		
8.4 Duración del tratamiento	No ofrece información sobre la duración de toda la intervención	El tiempo de duración de toda la intervención es menor a 4 semanas (inclusive)	El tiempo de duración de toda la intervención es entre 5 y 8 semanas (ambos inclusive)	El tiempo de duración de toda la intervención es entre 9 y 12 semanas (ambos inclusive)	El tiempo de duración de toda la intervención es mayor de 17 semanas (inclusive)		

## VALIDEZ INTERNA (continuación)

9. INSTRUMENTOS DE MEDIDA					
Indicador	0		0,5	1	
<b>9.1</b> Momento medición de la VD	No se menciona si los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y experimental	Más de un instrumento de medida o el único no se ha aplicado al mismo tiempo en el grupo control y experimental	Al menos uno de todos los instrumentos de medida no se ha aplicado al mismo tiempo en el grupo control y experimental	Todos los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y experimental	
	0		1		
<b>9.2</b> Instrumentos de medida (homogeneidad )	No queda claro si los instrumentos de medida que se han aplicado tanto en el grupo control como experimental son iguales o diferentes	Más de un instrumento de medida o el único no es el mismo para el grupo control y experimental	Al menos uno de todos los instrumentos de medida que han sido aplicados tanto en el grupo experimental como control es diferente	Todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales	
	0		0,25	0,5	1
<b>9.3</b> Seguimiento de la VD	Después de finalizada la intervención no se mide nuevamente la variable dependiente	Entre 3 y 5 meses (ambos incluidos) después de finalizada la intervención se mide nuevamente la variable dependiente	Entre 6 y 11 meses (ambos incluidos) después de finalizada la intervención se mide nuevamente la variable dependiente	Un año después de finalizada la intervención se mide nuevamente la variable dependiente	
10. SESGOS DEL INVESTIGADOR Y DE LOS SUJETOS					
Indicador	0		0,5	1	
<b>10.1</b> Sesgo de procedimiento	No aparece información sobre la persona que lleva a cabo la intervención en el grupo control y el grupo experimental	No es la misma persona quien lleva a cabo la intervención tanto en el grupo control como experimental	No es la misma persona quien lleva a cabo la intervención en el grupo control y experimental pero se controlan las posibles diferencias a través de formación	Es la misma persona quien lleva a cabo la intervención tanto en el grupo control como experimental	
	0		1		
<b>10.2</b> Enmascaramiento del sujeto	No ofrece información si los sujetos conocen o no si han sido asignados al grupo control o experimental	Los sujetos conocen si han sido asignados al grupo control o experimental	Los sujetos desconocen si forman parte del grupo control o experimental		
<b>10.3</b> Enmascaramiento del evaluador	No ofrece información si los evaluadores conocen o no los objetivos de la investigación	Los evaluadores conocen los objetivos de la investigación	Los evaluadores desconocen los objetivos de la investigación		
<b>10.4</b> Incentivos/recompensas	No ofrece información si los sujetos reciben o	Los sujetos reciben recompensas por su	Los sujetos participan en la investigación de manera voluntaria sin		

	no recompensas por su participación en el estudio	participación en la investigación	recibir recompensas	
	<b>TOTAL VALIDEZ INTERNA</b>			

**VALIDEZ EXTERNA**

<b>11. INTERPRETACIÓN RESULTADOS</b>				
<b>Indicador</b>	<b>0</b>	<b>0,5</b>	<b>1</b>	<b>Valor</b>
<b>11.1</b> Discusión causas–intervención	La discusión de las causas por las cuales la intervención ha sido efectiva o no es muy débil	Se discuten las causas por las cuales la intervención ha sido efectiva o no dejando fuera aspectos relevantes a considerar	Se discuten las causas por las cuales la intervención ha sido efectiva o no	
<b>11.2</b> Discusión intervención–teoría	La discusión de los resultados de la intervención en función de la teoría y estudios empíricos es muy débil	Los resultados de la intervención se discuten en función de la teoría y estudios empíricos quedando fuera aspectos relevantes a considerar	Los resultados de la intervención se discuten en función de la teoría y estudios empíricos relevantes	
<b>11.3</b> Estrategia de análisis–intervención	No informa sobre la estrategias de análisis que se han utilizado para los sujetos que no completaron la intervención	La estrategia de análisis aplicada no ha sido por intención de tratar ( <i>intention to treat analysis</i> ) pero describe aquella que se ha utilizado para quienes no terminaron el tratamiento	La estrategia de análisis aplicada ha sido por intención de tratar ( <i>intention to treat analysis</i> )	
	<b>0</b>	<b>1</b>		
<b>11.4</b> Limitaciones intervención	No informa de las debilidades de la intervención	Informa de las dificultades para la implementación de la intervención		
<b>11.5</b> Futuras líneas de investigación	A partir de los hallazgos encontrados no se sugieren posibles investigaciones futuras	A partir de los hallazgos encontrados se sugieren posibles investigaciones futuras		
<b>12. GENERALIZACIÓN RESULTADOS</b>				
<b>Indicador</b>	<b>0</b>	<b>1</b>		
<b>12.1</b> Interacción selección–tratamiento (validez de población)	Se generalizan los resultados de la intervención a una población que no ha sido estudiada	Se generalizan los resultados de la intervención solo a la población que ha sido estudiada		
<b>12.2</b> Interacción contexto–tratamiento (validez ecológica)	Se generalizan los resultados de la intervención a una situación o contexto que no ha sido estudiada	Se generalizan los resultados de la intervención a situaciones o contextos que han sido estudiados		
<b>TOTAL VALIDEZ EXTERNA</b>				

**VALIDEZ DE CONCLUSIÓN ESTADÍSTICA**

<b>13. CONCLUSIÓN ESTADÍSTICA</b>					
<b>Indicador</b>	<b>0</b>		<b>0,5</b>	<b>1</b>	
<b>13.1</b> Tamaño muestra	El tamaño de la muestra tanto en el grupo control como experimental es menor de 10 sujetos		El tamaño de la muestra tanto en el grupo control como experimental es entre 10 y 30 sujetos (ambos valores inclusive)	El tamaño de la muestra tanto en el grupo control como experimental es mayor de 30 sujetos	
<b>13.2</b> Fiabilidad instrumentos de medición	Los instrumentos de medición no han sido validados	No hace mención a la fiabilidad de los instrumentos	Al menos la fiabilidad de uno de todos los instrumentos de medición es menor a 0,80	La fiabilidad de todos los instrumentos de medición es mayor a 0,80	
	<b>0</b>		<b>1</b>		
<b>13.3</b> Supuestos del modelo estadístico	Las pruebas estadísticas no tienen en cuenta los supuestos que deben cumplir los datos		Las pruebas estadísticas aplicadas tienen en cuenta los supuestos que deben cumplir los datos		
<b>13.4</b> Magnitud del efecto	No ofrece información con respecto a la magnitud del efecto		Ofrece información de la magnitud del efecto		
	<b>0</b>	<b>0,25</b>	<b>0,5</b>	<b>1</b>	
<b>13.5</b> Nivel de confianza	No hace mención al nivel de confianza	El nivel de confianza es de 90%	El nivel de confianza es de 95%	El nivel de confianza es de 99%	
<b>TOTAL VALIDEZ DE CONCLUSIÓN ESTADÍSTICA</b>					

**Anexo 17. Estudios seleccionados al azar para la fiabilidad inter-codificadores**

Estudios	Nº
Giraud, G. (1997)	1
AbuSeileek, A. (2007)	2
O'Donnell, A., Larson, C., Dansereau, D. & Rocklin, T. (1986)	3
O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. & Young, M. (1986)	4
Courtney, D., Courtney, M. & Nicholson, C. (1992)	5
Hassan, M. Fong, S. & Idrus, R. ( 2011)	6
Klein J. & Pridemore, D. (1992)	7
Cheng, Y. & Ku, H. (2009)	8
Kiliç, D. (2008)	9
Koç, Y., Doymuş, K., Karaçöp, A. & Şimşek, U. (2010)	10
Borresen, C. (1990)	11
Crooks, S., Klein, J., Savenye, W. & Leader, L. (1998)	12
Nihalani, P., Mayrath, M. & Robinson, D. (2011)	13
Fantuzzo J., Riggio, R. E., Connelly, S. & Dimeff, L. A (1989)	14
Ganter, S. (1994)	15
Doymuş K., Şimşek, U. & Karaçöp, A. (2009)	16
Yaeger, P., Marra, R., Costanzo, F. & Gray, G. (1999)	17
Oh, H. (1988)	18
Reglin, G. (1990)	19
Valentino, V. (1988)	20
Numpraset, W. (2006)	21
Doymuş, K. (2008a)	22
Doymuş, K., Karacop, A. & Simsek, U. (2010)	23
Turner, J. (1995)	24
Pavese, A. (1993)	25
Lovelace, T. & McKnight, C. (1980)	26

## Anexo 18. Resultado índice de Kappa

Tabla de contingencia calidad1 \* calidad2

			calidad2		Total
			Baja	Media	
calidad1	Baja	Recuento	2	0	2
		% del total	8,3%	0,0%	8,3%
	Media	Recuento	0	19	19
		% del total	0,0%	79,2%	79,2%
	Alta	Recuento	0	3	3
		% del total	0,0%	12,5%	12,5%
	Total	Recuento	2	22	24
		% del total	8,3%	91,7%	100,0%

Medidas simétricas

		Valor	Error típ. asint. <sup>a</sup>	T aproximada <sup>b</sup>	Sig. aproximada
Medida de acuerdo	Kappa	,532	,221	4,130	,000
N de casos válidos		24			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.



**Anexo 19. Versión final de la escala de calidad metodológica según fiabilidad inter-observadores**

**Código:**

**Fecha:**

**Evaluator:**

Información general documento primario	
Nombre los autores	
Título del documento	
Año	

Criterios de inclusión	Si	No
VI: programa aprendizaje cooperativo y aprendizaje individual (grupo control inactivo)		
VD: rendimiento ( <i>outcome</i> )		
Estudiantes universitarios de todos los niveles		
Al menos un grupo control y experimental		
Con diseño pre-test y pos-test o solo post-test		
Suficiente información para el cálculo de la magnitud del efecto (ME)		

**PUNTUACIÓN POR TIPO DE VALIDEZ**

Tipos de validez	Valor total	%	Total
Validez de constructo	MAX 13	30,95%	
Validez interna	MAX 17	40,47%	
Validez externa	MAX 7	16,66%	
Validez de conclusión estadística	MAX 5	11,90%	
<b>Total</b>	<b>42</b>	<b>100%</b>	

**CALIDAD METODOLÓGICA**

Niveles de calidad	Valor
Bajo (entre 0 y 4, 99 ambos inclusive)	
Medio (entre 5 y 6,99 ambos inclusive)	
Alto (entre 7 y 10, ambos inclusive)	

## VALIDEZ DE CONSTRUCTO

1. CONSIDERACIONES TEÓRICAS, OBJETIVOS E HIPÓTESIS DEL ESTUDIO					
Indicador	0		0,5	1	Valor
1.1 Teoría	El diseño de la intervención se sustenta en un marco teórico débil que refleja un bajo dominio en el área de estudio dejando afuera una gran cantidad de aspectos relevantes de la teoría		El diseño de la intervención se sustenta en un marco teórico que refleja conocimiento del área de estudio pero deja de mencionar algunos aspectos relevantes de la teoría	El diseño de la intervención se sustenta en un marco teórico claro y bien articulado	
1.2 Estudios previos	No hace mención a estudios previos relacionados con la intervención		Se mencionan estudios previos los cuales se han implementado en contextos no similares en donde la intervención ha sido eficaz	Se mencionan estudios previos los cuales se han implementado en contextos similares donde la intervención ha sido eficaz	
	0			1	
1.3 Objetivos e hipótesis	Los objetivos e hipótesis del estudio no son claros			Los objetivos e hipótesis del estudio son claros	
2. DEFINICIÓN DE CONSTRUCTOS					
Indicador	0			1	
2.1 Definición teórica VI	La variable independiente no está claramente definida a nivel teórico			La variable independiente está claramente definida a nivel teórico	
2.2 Definición operacional al VI	La variable independiente no está claramente definida a nivel operacional porque describe solo el grupo de intervención sin explicar cómo difiere del grupo control			La variable independiente está claramente definida a nivel operacional porque describe con detalle cómo el grupo de intervención difiere del grupo control	
2.3 Definición teórica VD	La variable dependiente no está claramente definida a nivel teórico			La variable dependiente está claramente definida a nivel teórico	
	0		0,5	1	
2.4 Adecuación instrumentos de medida VD	No se describe más de un instrumento de medida o el único que se ha aplicado para medir la variable dependiente		Se describe el único o al menos uno de todos los instrumentos que han sido aplicados para medir la variable dependiente	Se describen todos o el único instrumento de medición que se han aplicado para medir la variable dependiente	
	0		0,5	1	
2.5 Validación instrumentos de medida VD	No hace mención si los instrumentos que se han aplicado para medir la variable dependiente han sido o no validados/estandarizados	Más de un instrumento de medida o el único que se ha aplicado para medir la variable dependiente no ha sido validado/estandarizados	Al menos uno de todos los instrumentos que han sido aplicados para medir la variable dependiente no ha sido validado/estandarizados	El único o todos los instrumentos de medición que se han aplicado para medir la variable dependiente han sido validados/estandarizados	

## VALIDEZ DE CONSTRUCTO (continuación)

3. SESGOS OPERACIONALIZACIÓN CAUSA-EFECTO/REACTIVIDAD SUJETOS				
Indicador	0		1	
<b>3.1</b> Sesgo único método medición VD	La variable dependiente es medida solo a través de un mismo método		La variable dependiente es medida con más de un método	
<b>3.2</b> Interacción administración de pruebas y tratamiento	Los instrumentos de medida dan ventaja al grupo experimental		Los instrumentos de medida no dan ventaja al grupo experimental	
	0		1	
<b>3.3</b> Difusión/imitación del tratamiento	No ofrece información con respecto a si los participantes del grupo control y experimental comparten o no información o intentan copiar el tratamiento	Los sujetos del grupo control y experimental comparten información o intentan copiar el tratamiento	Se garantiza que los sujetos del grupo control y experimental no comparten información o intentan copiar el tratamiento	
<b>3.4</b> Interacción de tratamientos intra-sujetos	No ofrece información que demuestre que ninguno de los sujetos del grupo control ha participado posteriormente en el grupo experimental	Sujetos del grupo control han participado posteriormente en el grupo experimental (contaminación <i>-cross-over-</i> )	Ningún sujeto del grupo control ha participado posteriormente en el grupo experimental (contaminación <i>cross-over</i> )	
<b>3.5</b> Rivalidad compensatoria de los sujetos	No ofrece información sobre las técnicas que se utilizan para evitar que el grupo control intente competir con el grupo experimental	No se compensa al grupo control con técnicas cualitativas para evitar que busquen competir con el grupo experimental	Se compensa al grupo control con técnicas cualitativas (como entrevistas no estructuradas y observación directa) para evitar que busquen competir con el grupo experimental	
TOTAL VALIDEZ DE CONSTRUCTO				

## VALIDEZ INTERNA

4. DISEÑO INVESTIGACIÓN					
Indicador	0		0.5	1	
4.1 Situación basal (homogeneidad)	No ofrece información con respecto a si el grupo control y experimental son inicialmente equivalentes en cuanto a la variable dependiente	Menciona que el grupo control y experimental no son inicialmente equivalentes en cuanto a la variable dependiente	Antes de la intervención, se describen las características demográficas de los participantes en el grupo control y experimental como evidencia de que ambos son inicialmente equivalentes	Antes de la intervención, se demuestra estadísticamente que el grupo control y experimental son inicialmente equivalentes en cuanto a la variable dependiente	
5. CRITERIOS INCLUSIÓN/EXCLUSIÓN MUESTRA					
Indicador	0		1		
5.1 Características muestra	No menciona las características de la muestra		Menciona las características de la muestra		
5.2 Inclusión-exclusión de sujetos	Antes de la intervención, no explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos		Antes de la intervención, explica las decisiones por las cuales algunos sujetos del estudio formaron parte de la investigación y otros quedaron excluidos	Antes de la intervención, todos los sujetos del estudio han formado parte de la investigación, no se han excluido sujetos	
6. PROCEDIMIENTO SELECCIÓN/ASIGNACIÓN MUESTRA					
Indicador	0		1		
6.1 Selección grupos (población)	No ofrece información sobre cómo han sido seleccionados los grupos (población)	No se han seleccionado aleatoriamente los grupos que van a participar en el estudio	Se han seleccionado aleatoriamente los grupos que van a participar en el estudio		
6.2 Asignación grupos	No ofrece información sobre cómo han sido asignados el grupo control y el grupo experimental	Los grupos no han sido asignados aleatoriamente como grupo control o grupo experimental	Los grupos han sido asignados aleatoriamente como grupo control o grupo experimental		
	0		0.5	1	
6.3 Asignación sujetos	No ofrece información sobre cómo han sido asignados los sujetos al grupo control y al grupo experimental	Los sujetos no han sido asignados aleatoriamente al grupo control y experimental	Se ha aplicado algún procedimiento de control de variables extrañas como por ejemplo: emparejamiento, bloqueo o estratificación	Los sujetos han sido asignados aleatoriamente al grupo control y experimental	

## VALIDEZ INTERNA (continuación)

7. MORTALIDAD						
Indicador	0		0,5		1	
7.1 Mortalidad (sujetos del estudio)	El porcentaje de participantes que no ha llegado a terminar el estudio supera el 20% de la muestra original		El porcentaje de participantes que no ha llegado a terminar el estudio es >0% y <20% de la muestra original (ambos inclusive)		El porcentaje de participantes que ha iniciado el estudio se mantiene al terminar la investigación, no hay pérdida de sujetos	
	0		0,25	0,5	1	
7.2 Mortalidad (grupo control y experimental)	No ofrece información de la diferencia de pérdidas de participantes entre el grupo control y experimental	El porcentaje de la diferencia de pérdidas de participantes entre el grupo control y experimental es mayor del $\geq 20\%$	El porcentaje de la diferencia de pérdida de participantes entre el grupo control y experimental es >0% y <20%	El porcentaje de la diferencia de pérdida de participantes entre el grupo control y experimental es igual	El porcentaje de sujetos en el grupo control y experimental se mantiene al terminar la investigación	
8. ADMINISTRACIÓN DEL TRATAMIENTO						
Indicador	0			1		Valor
8.1 Persona que administra el tratamiento	No ofrece información sobre quién es la persona que lleva a cabo la intervención			Ofrece información sobre quién es la persona que lleva a cabo la intervención		
	0		0,5	1		
8.2 Lugar del tratamiento	No ofrece información sobre el lugar en donde se ha llevado a cabo la intervención		La intervención se ha llevado a cabo en laboratorios o aulas especialmente diseñadas para el estudio	La intervención se ha llevado a cabo en aulas ordinarias y no en laboratorios o aulas especialmente diseñadas para el estudio		
8.3 Distribución duración tratamiento	No ofrece información del número, frecuencia y duración de cada sesión	Ofrece solo información del número, frecuencia o duración de cada sesión	Ofrece información en dos de los aspectos relacionados con la duración de la intervención: número, frecuencia o duración de cada sesión	Ofrece información sobre la intervención en cuanto a número, frecuencia y duración de cada sesión		

**VALIDEZ INTERNA (continuación)**

<b>9. INSTRUMENTOS DE MEDIDA</b>					
<b>Indicador</b>	<b>0</b>		<b>0,5</b>	<b>1</b>	
<b>9.1</b> Momento medición de la VD	No se menciona si los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y experimental	Más de un instrumento de medida o el único no se ha aplicado al mismo tiempo en el grupo control y experimental	Al menos uno de todos los instrumentos de medida no se ha aplicado al mismo tiempo en el grupo control y experimental	Todos los instrumentos de medida se han aplicado al mismo tiempo en el grupo control y experimental	
	<b>0</b>			<b>1</b>	
<b>9.2</b> Instrumentos de medida (homogeneidad)	No se menciona si los instrumentos de medida que se han aplicado tanto en el grupo control como experimental son iguales o diferentes	El único o más de un instrumento de medida no es el mismo para el grupo control y experimental	El único o al menos uno de todos los instrumentos de medida que han sido aplicados tanto en el grupo experimental como control es diferente	El único o todos los instrumentos de medida que han sido aplicados en el grupo control y experimental son iguales	
<b>10. SEGOS DEL INVESTIGADOR Y DE LOS SUJETOS</b>					
<b>Indicador</b>	<b>0</b>		<b>0,5</b>	<b>1</b>	
<b>10.1</b> Sesgo de procedimiento	No aparece información sobre la persona que lleva a cabo la intervención en el grupo control y el grupo experimental	No es la misma persona quien lleva a cabo la intervención tanto en el grupo control como experimental	No es la misma persona quien lleva a cabo la intervención en el grupo control y experimental pero se controlan las posibles diferencias a través de formación	Es la misma persona quien lleva a cabo la intervención tanto en el grupo control como experimental	
	<b>0</b>			<b>1</b>	
<b>10.2</b> Enmascaramiento del sujeto	No ofrece información si los sujetos conocen o no si han sido asignados al grupo control o experimental	Los sujetos conocen si han sido asignados al grupo control o experimental	Los sujetos desconocen si forman parte del grupo control o experimental		
<b>10.3</b> Enmascaramiento del evaluador	No ofrece información si los evaluadores conocen o no los objetivos de la investigación	Los evaluadores conocen los objetivos de la investigación	Los evaluadores desconocen los objetivos de la investigación		
<b>10.4</b> Incentivos / recompensas	No ofrece información si los sujetos reciben o no recompensas por su participación en el estudio	Los sujetos reciben recompensas por su participación en la investigación	Los sujetos participan en la investigación de manera voluntaria sin recibir recompensas		
	<b>TOTAL VALIDEZ INTERNA</b>				

## VALIDEZ EXTERNA

11. INTERPRETACIÓN RESULTADOS				
Indicador	0	0,5	1	Valor
11.1 Discusión causas-intervención	La discusión de las causas por las cuales la intervención ha sido efectiva o no es muy débil	Se discuten las causas por las cuales la intervención ha sido efectiva o no dejando fuera aspectos relevantes a considerar	Se discuten las causas por las cuales la intervención ha sido efectiva o no	
11.2 Discusión intervención-teoría	La discusión de los resultados de la intervención en función de la teoría y estudios empíricos es muy débil	Los resultados de la intervención se discuten en función de la teoría y estudios empíricos quedando fuera aspectos relevantes a considerar	Los resultados de la intervención se discuten en función de la teoría y estudios empíricos relevantes	
11.3 Estrategia de análisis-intervención	No informa sobre la estrategias de análisis que se ha utilizado para los sujetos que no completaron la intervención	La estrategia de análisis aplicada no ha sido por intención de tratar ( <i>intention to treat analysis</i> ) pero describe aquella que se ha utilizado para quienes no terminaron el tratamiento	La estrategia de análisis aplicada para los sujetos que han complementado o no la intervención ha sido por intención de tratar ( <i>intention to treat analysis</i> )	
	0	1		
11.4 Limitaciones intervención	No informa de las debilidades del estudio	Informa de las debilidades del estudio		
11.5 Futuras líneas de investigación	A partir de los hallazgos encontrados no se sugieren posibles investigaciones futuras	A partir de los hallazgos encontrados se sugieren posibles investigaciones futuras		
12. GENERALIZACIÓN RESULTADOS				
Indicador	0	1		
12.1 Interacción selección-tratamiento o (validez de población)	Se generalizan los resultados de la intervención a una población que no ha sido estudiada	Se generalizan los resultados de la intervención solo a la población que ha sido estudiada		
12.2 Interacción contexto-tratamiento o (validez ecológica)	Se generalizan los resultados de la intervención a una situación o contexto que no ha sido estudiada	Se generalizan los resultados de la intervención a situaciones o contextos que han sido estudiados		
TOTAL VALIDEZ EXTERNA				

## VALIDEZ DE CONCLUSIÓN ESTADÍSTICA

13. CONCLUSIÓN ESTADÍSTICA						
Indicador	0		0.25	0.50	0.75	1
13.1 Tamaño muestra	El tamaño de la muestra tanto en el grupo control como experimental es menor de 10 sujetos		Al menos el tamaño de la muestra o bien del grupo experimental o bien del grupo control es entre 10 y 30 sujetos o menor de 10	El tamaño de la muestra tanto en el grupo control como experimental es entre 10 y 30 sujetos (ambos valores inclusive)	Al menos el tamaño de la muestra o bien del grupo experimental o bien del grupo control es entre 10 y 30 sujetos o mayor de 30	El tamaño de la muestra tanto en el grupo control como experimental es mayor de 30 sujetos
	0		0.5		1	
13.2 Fiabilidad instrumentos de medición	No hace mención de la fiabilidad de los instrumentos de medida	Más de un instrumento de medida o el único no han sido validados	Al menos la fiabilidad del único o de uno de todos los instrumentos de medición es menor a 0.80		La fiabilidad del único o de todos los instrumentos de medición es mayor a 0.80	
	0			1		
13.3 Supuestos del modelo estadístico	Las pruebas estadísticas no tienen en cuenta los supuestos que deben cumplir los datos			Las pruebas estadísticas aplicadas tienen en cuenta los supuestos que deben cumplir los datos		
13.4 Magnitud del efecto	No ofrece información con respecto a la magnitud del efecto			Ofrece información de la magnitud del efecto		
	0	0.25		0.5	1	
13.5 Nivel de confianza	No hace mención al nivel de confianza	El nivel de confianza es de 90%		El nivel de confianza es de 95%	El nivel de confianza es de 99%	
TOTAL VALIDEZ DE CONCLUSIÓN ESTADÍSTICA						

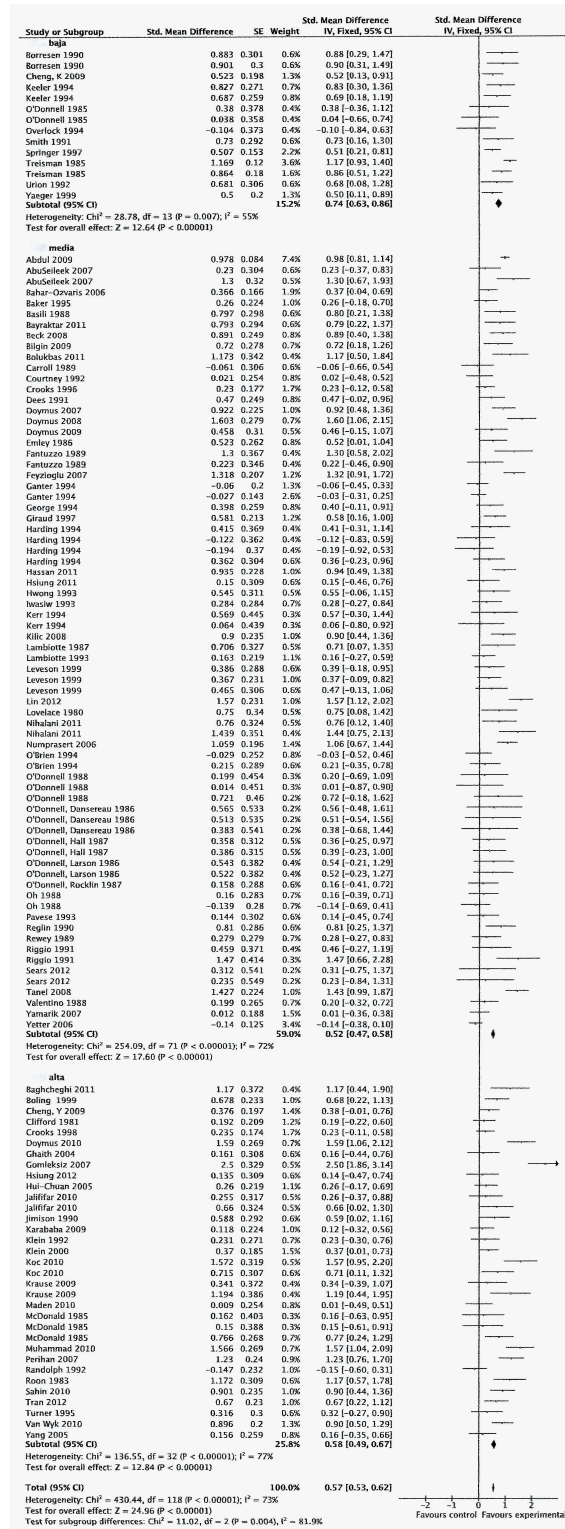


**Anexo 20. Número de resultados por estudios primarios**

<i>Outcome s</i>	Estudio	Año publicación
1	Lovelace, T. & McKnight, C.	1980
1	Emley, W.	1986
1	Basili, P.	1988
1	Valentino, V.	1988
1	Carroll, E.	1989
1	Jimison, L.	1990
1	Reglin, G.	1990
1	Dees, R.	1991
1	Smith, M., Hinckley, C. & Volk, G.	1991
1	Courtney, D., Courtney, M. & Nicholson, C.	1992
1	Klein, J. & Pridemore, D.	1992
1	Randolph, W.	1992
1	Urion, D. & Davidson, N.	1992
1	Iwasiw, C. & Goldenberg, D.	1993
1	Lambiotte, J., Skaggs, L. & Dansereau, D.	1993
1	Hwong, N., Caswell, A., Johnson, D. & Johnson, R.	1993
1	Pavese, A.	1993
1	George, P.	1994
1	Overlock, T.	1994
1	Baker, L.	1995
1	Turner, J.	1995
1	Crooks, S., Klein, J., Jones, E. & Dwyer, H.	1996
1	Giraud, G.	1997
1	Springer, L.	1997
1	Crooks, S., Klein, J., Savenye, W. & Leader, L.	1998
1	Yaeger, P., Marra, R., Costanzo, F. & Gray, G.	1999
1	Boling, N. & Robinson, D.	1999
1	Klein, J. & Schnackenberg, H.	2000
1	Ghaith, G. & El-malak, M.	2004
1	Yang, A.	2005
1	Hui-Chuan, L.	2005
1	Bahar-Özvaris, S., Çuhadaroglu, F., Turan, S. & Peters, A.	2006
1	Numprasert, W.	2006
1	Yetter, G., Gutkin, T., Saunders, A., Galloway, A., Sobansky, R. & Song, S.	2006
1	Yamarik, S.	2007
1	Feyzioğlu, B., Akçay, H. & Şahin-Pekmez, E.	2007
1	Doymuş, K.	2007
1	Gömleksiz, M.	2007
1	Perihan, A. & Kamuran, T.	2007
1	Beck, L. & Chizhik, A.	2008
1	Doymuş, K.	2008a
1	Kiliç, D.	2008
1	Tanel, Z. & Erol, M.	2008
1	Cheng, K.	2009
1	Abdul, S.	2009
1	Bilgin, I.	2009
1	Karababa, C.	2009
1	Cheng, Y. & Ku, H.	2009
1	Doymuş, K., Şimşek, U. & Karaçöp, A.	2009
1	Doymuş, K., Karaçöp, A. & Şimşek, U.	2010
1	Maden, S.	2010
1	Muhammad, Z.	2010

1	Şahin, A.	2010
1	Van Wyk, M.	2010
1	Hassan, M., Fong, S. & Idrus, R.	2011
1	Baghcheghi, N., Reza, H. & Rezaei, K.	2011
1	Bayraktar, G.	2011
1	Bölükbaş, F., Keskin, F. & Polat, M.	2011
1	Hsiung, C.	2011
1	Lin, Z.	2012
1	Tran, V. & Lewis, R.	2012
1	Hsiung, C.	2012
1	O'Donnell, A., Dansereau, D., Rocklin, T., Larson, C., Hythecker, V., Young, M. & Lambiotte, J.	1987
1	Lambiotte, J., Dansereau, D., Rocklin, T., Fletcher, B., Hythecker, V., Larson, C. & O'Donnell, A.	1987
1	Roon, R., Van Pilsum, J., Harris, I., Rosenberg, P., Johnson, R., Liaw, C. & Rosenthal, L.	1983
1	Rewey, K., Dansereau, D., Skaggs, L., Hall, R. & Pitre, U.	1989
1	Clifford, J.	1981
2	O'Donnell, A., Dansereau, D., Rocklin, T., Hythecker, V., Lambiotte, J., Larson, C. & Young, M.	1985
2	Treisman, P.	1985
2	Oh, H.	1988
2	Fantuzzo, J., Riggio, R., Connelly, S. & Dimeff, L.	1989
2	Borresen, C.	1990
2	Riggio, R., Fantuzzo, J., Connelly, S. & Dimeff, L.	1991
2	Ganter, S.	1994
2	Keeler, C. & Steinhorst, R.	1994
2	O'Brien, G. & Peters, J.	1994
2	Kerr, D. & Murthy, U.	1994
2	AbuSeileek, A.	2007
2	Krause, U., Stark, R. & Mandl, H.	2009
2	Jalilifar, A.	2010
2	Koç, Y., Doymuş, K., Karaçöp, A. & Şimşek, U.	2010
2	Nihalani, P., Mayrath, M. & Robinson, D.	2011
2	Sears, D. & Pai, H.	2012
2	O'Donnell, A., Larson, C., Dansereau, D. & Rocklin, T.	1986
2	O'Donnell, A., Dansereau, D., Hall, R. & Rocklin, T.	1987
3	McDonald, B., Larson, C., Dansereau, D. & Spurlin, J.	1985
3	Leveson, L.	1999
3	O'Donnell, A., Dansereau, D., Hythecker, V., Hall, R., Skaggs, I., Lambiotte, J., Young, M.	1988
3	O'Donnell, A., Dansereau, D., Hythecker, V., Larson, C., Rocklin, T., Lambiotte, J. & Young, M.	1986
4	Harding, R. & Fletcher, R.	1994
<b>119</b>	<b>TOTAL</b>	

## Anexo 21. Forest plot calidad metodológica estudios primarios



**Anexo 22. Resumen características principales de los estudios incluidos en el meta-análisis**

Variables extrínsecas				Variables ecológicas										Variables metodológicas							Síntesis cuantitativa																					
	Tiempo de fuente	Tiempo de documento	Lugar de procedencia	Tamaño de la muestra inicial	Asignaturas ***	Método AC	Duración programa	Nº integrantes	Grupos pares/impares	Grupos homogéneos/heterogéneos	Estructura de recompensa	Estructura de tarea	Igualdad oportunidades puntuación	Sesgos investigador	Asignación	Situación basal	Instrumentos de medida	Calidad metodológica	Nº sujetos del grupo control/ grupo experimental*	Outcomes	Magnitud del efecto subgrupo/medida	SE	Magnitud del efecto global**	SE																		
Estudio																																										
1. Abdul, S. (2009)	3	4	1	645 student teachers from three teacher training colleges from University Pendidikan Sultan Idris (UPSI) who were enrolled in the introductory statistics courses	Estadística	Cooperative Learning	2	2-3	4	3				1	0	1	2	2	Cooperative learning group, Ne=370; lecture method, Ne=275	Classroom test	0,978		0,084																			
2. AbuSelleek, A. (2007)	3	4	1	130 freshmen in the BA program of English Language and Literature, College of Arts at King Saud University	Lengua Extranjera	Cooperative Computer-Assisted language learning (CALL) (a distance)	4	2	1	3				1	0	1	1	2		Cooperative computer-mediated communication technique (CoCMC), Ne=36; collective traditional technique, Ne=34*	Listening	1,163	0,315	1,300	0,320																	
																					Speaking	1,456	0,326																			
						Cooperative Learning															2	1	3			1	0	1	1	2		Cooperative traditional technique, Ne=36; collective traditional technique, Ne=34*	Listening	0,572	0,309	0,230	0,304					
																																	Speaking	-0,090	0,304							
3. Baghchehghi, N., Reza, H. & Rezaei, K. (2011)	3	4	5	34 nursing students in their first year of training, Arak University of Medical Sciences in Iran	Enfermería	Cooperative Learning	6	4	1	2				1	1	1	1	3	Cooperative learning, Ne=16; traditional learning, Ne=18	Change score of communication skill	1,170		0,372																			

4. Bahar-Oyaris, S., Çuhadroğlu, F., Turan, S. & Peters, A. (2006)	3	4	5	150 fifth-year students in the two-week Adolescent Development and Psychopathology Course within the sixweek Mental Health Clerkship	Psicología	Student Team Learning (STL) + Problem-Based Learning (PBL)	2	8-10	4	3									Experimental group, Ne=67; control group, Ne=83	Gains achievement test		0,366	0,166	
5. Baker, L. (1995)	1	2	5	81 undergraduates in Introductory Computer Science Programming	Tecnología	Cooperative study group in an introductory computer science programming (face to face)	4	4-5	4	2									Cooperative study group method, Ne=35; traditional individual learning method, Ne=33	Gains achievement score	0,260	0,224		
6. Basili, P. (1988)	1	2	5	62 suburban community college students	Química	Cooperative Learning	2	3-4-5	4	2									Cooperative groups, Ne=35; control group, Ne=27	Misconceptions Matter	0,812	0,299	0,797	0,298
																				Misconceptions Energy	0,681	0,295		
																				Misconceptions Gases	0,569	0,303		
																				Misconceptions Liquidos	0,891	0,312		
																				Misconceptions Solids	1,069	0,318		
7. Bayraktar, G. (2011)	3	4	2	50 university students who were enrolled in two classes of the general gymnastics course in the first year of Physical Education and Athletics Teaching	Educación Física	Learning Together	2	5	2	2								Cooperative group, Ne=25; traditional group, Ne=25	Gains in gymnastics academic achievement test	0,774	0,293	0,793	0,294	
8. Beck, L. & Chizhik, A. (2008)	4	1	5	71 college students at our educational experiment involving two different sections of a Computer Science I course	Tecnología	Cooperative learning techniques in computer science courses (face to face)	6	No appearance	3	3									Experimental group, Ne=34; control group, Ne=37	Final exam		0,891	0,249	
																				Gains in gymnastics practice and evaluation form	0,812	0,294		
9. Bilgin, I. (2009)	3	4	2	55 first year university students from two intact classes of a Chemistry Course	Química	Cooperative Learning	1	2,4	1	2									Experimental group, Ne=28; control group, Ne=27	Gains Acid and Bases Achievement Test ABAT		0,720	0,278	
10. Boling, N. & Robinson, D. (1999)	3	4	5	115 undergraduates enrolled in five classes offered in the Department of Health, Physical Education, Recreation, and Sport at Mississippi State University	Comprehsion Lectora	Informal Cooperative Learning	7	4	1	3									Cooperative learning group, Ne=39; individual learning group, Ne=39	Gain score		0,678	0,233	
11. Bolukbas, F., Keskin, F. & Polat, M. (2011)	3	4	2	40 students who learn Turkish as a second language at Istanbul University Language Center	Lengua Extranjera	Ask together learn together (tecnic de Acikgöz)	2	3-4	4	2									Cooperative group, Ne=20; traditional group, Ne=20	Gains in Reading Comprehensi on Skills Achievement Test		1,173	0,342	

12. Borresen, C. (1990)	1	4	5	106 college students in six Introductory Statistics classes	Estadística	Cooperative Learning	6	3-4-5	4	3					3	0	0	2	1	Cooperative assigned groups, N=34; traditional method, N=37*	Means adjusted Final point	0,883	0,301
13. Carroll E. (1989)	3	2	5	54 college enrolled in two sections of GENAD 3113, Written Business Communication at Oklahoma State University	Lenguaje Escrito	Cooperative Learning	4	5	2	3					1	2	1	1	2	Cooperative voluntary groups, N=35; traditional method, N=37*	Means adjusted Final point	0,901	0,300
14. Cheng, K. (2009)	3	4	1	109 college students at year 1 technical school who attended two accounting classes in Taiwan	Administración	STAD, web-based collaborative learning instruction (face to face)	3	No aparece	3	3					1	0	1	1	1	Cooperative method, N=32; traditional method, N=16	Gains 75-items multiple choice test	0,061	0,306
15. Cheng, Y. & Ku, H. (2009)	3	4	1,5	105 undergraduate students who took a Technology in Education course at a university in the western United States	Tecnología	Reciprocal Peer Tutoring, educational technology course (face to face)	3	3-4	4	3					1	2	1	2	3	Reciprocal Peer Tutoring (RPT), N=56; non-RPT, N=49	Total points on Achievement score	0,376	0,197
16. Clifford, J. (1981)	1	4	5	92 freshmen at Queens College in New York	Lenguaje Escrito	Collaborative pedagogy/collaborative learning	4	6	1	3					2	1	1	2	3	Experimental group, N=43; control group, N=49	Gains Writing performance	0,192	0,209
17. Courtney, D., Courtney, M. & Nicholson, C. (1992)	4	1	5	62 graduate students in the school of education at a small state university	Estadística	Informal learning group	6	4	1	2					2	0	0	2	2	Cooperative learning, N=32; traditional method, N=30	Multiple-choice exam 1	0,021	0,254
18. Crooks, S., Klein, J., Jones, E. & Dwyer, H. (1996)	3	4	5	128 undergraduate education students from a large Southwestern University	Psicología	Cooperative Learning Computer-Based Instruction (CBI) (face to face)	6	2	1	3					3	1	0	1	2	Cooperative method, N=64; individual method, N=64	Practice-test performance	0,230	0,177
19. Crooks, S., Klein, J., Savenye, W. & Leader, L. (1998)	3	4	5	195 undergraduate students majoring in education and enrolled in an Educational Psychology course.	Psicología	Cooperative Learning Computer-Based Instruction (CBI) (face to face)	1	2	1	2					3	3	0	1	3	Cooperative learning, N=100; individual learning, N=50	Posttest performance	0,235	0,174
20. Dees, R. (1991)	1	4	5	72 students registered in a course designed to remove their deficiencies in high school algebra and geometry. Commuter	Matemáticas	Cooperative Learning	4	4-6	4	3					3	3	1	3	2	Cooperative group, N=32; individual group, N=40	Final exam total algebra	0,470	0,249

[illegible]

26. Fantuzzo, J., Riggio, R., Connelly, S. & Dineff, L. (1989)	3	4	5	100 undergraduate students enrolled in sections of Abnormal Psychology at California State University	Psicología	Reciprocal Peer Tutoring	6	2	1	3	3	3	1	1	1	2	Dyadic structured format, N=22; independent unstructured format, N=30*	Comprehensi ve examination	1,300	0,367
27. Feyzoğlu, B., Akay, H. & Şahin-Pekmez, E. (2007)	3	1	2	114 first year undergraduate students in the department of both chemistry and science education in Turkey (Term of 2005-2006 academic year)	Química	Computer Assisted Collaborative (face to face)	6	2	1	3	3	0	1	1	2	Computer assisted collaborative (CACL), N=54; individualistic learning (CAIL), N=60	Subgrupo Química Gains Achievement Scale (AS)	1,318	0,207	
28. Ganter, S. (1994)	1	4	5	100 students in first semester precalculus at the community college. Fall semester 1989-1990	Matemáticas	Professional Development Program (PDP) Workshop students	6	2	1	3	1	4	1	2	2	Workshop students, N=10; non-workshop matches, N=10	Posttest score	-	0,200	
29. George, P. (1994)	3	4	5	61 students in undergraduate Educational Psychology classes	Psicología	Think-pair-share	5	2	1	3	1	4	1	2	2	Workshop students, N=14; Non-workshop matches, N=14	Posttest score	-	0,143	
30. Ghaith, G. & El-malak, M. (2004)	3	4	1	48 students in an English course	Lengua Extranjera	Jigsaw II	2	4-5	4	3	1	1	2	1	3	Cooperative learning group, N=21; individual group, N=19	Gains test comprehensio n	0,161	0,308	
31. Giraud, G. (1997)	1	4	5	102 students in undergraduate Statistics course	Estadística	Cooperative Learning	6	5	2	3	1	0	1	2	2	Cooperative group, N=44; individual lecture, N=48	Test 2	0,581	0,213	
32. Gömleksiz, M. (2007)	3	4	2	66 engineering students participated in the study	Lengua Extranjera	Jigsaw II	4	5-6	4	3	1	1	3	1	3	Cooperative jigsaw II method, N=34; traditional teacher-centred teaching method, N=32	Gains Achievement test	2,500	0,329	
33. Harding, R. & Fletcher, R. (1994)	1	1	5	98 students in a remedial mathematics course RSD073, Tennessee Technological University	Matemáticas	STAD	6	4	1	2	1	2	1	3	0	1	1	2	0,415	0,369



																		Gains AAPPP math achievement examination		-0.194	0.370
																		Gains AAPPP math achievement examination		-0.122	0.362
																		Gains AAPPP math achievement examination		0.362	0.304
																		Gains Interpersonal Communicative Skills on ICS)		0.935	0.228
																		Academic achievement		0.150	0.309
																		Academic Achievement in Homework Tests-Unit1	0.314		
																		Academic Achievement in Homework Tests-Unit2	0.309		
																		Academic Achievement in Homework Tests-Unit3	0.313		
																		Academic Achievement in Homework Tests-Unit4	0.319		
																		Gains grammar test		0.260	0.219
																		Written exam	0.314		
																		Performance exam	0.308	0.545	0.311
																		Gains Cognitive Test	0.287	0.284	0.284
																		Pyscomotor test checklist	0.0073	0.283	

[illegible]

[illegible]

[illegible]

[illegible]

(1986)				Christian University		Scripted Cooperation		3	2	1	3	2	2	3	1	1	2	2	Passive trades Ne=16; individuals, Ne=14*	Passage 1	0.344	0.52 9	0.513 *	0.535
						Scripted Cooperation		2	1	1	3	2	2	3	1	1	2	2	Dyads Ne=17; individuals, Ne=14*	Passage 2	0.732	0.53 8		
						Scripted Cooperation														Passage 1	0.253	0.52 4	0.465 *	0.533
						Scripted Cooperation														Passage 2	0.947	0.54 2		
						Scripted Cooperation														Passage 1	0.852	0.39 1		
						Scripted Cooperation														Passage 2	-0.057	0.37 5	0.380	0.378
						Scripted Cooperation														Passage 1	0.247	0.35 9	0.038	0.358
						Scripted Cooperation														Passage 2	-0.169	0.35 8	0.038	0.358
						Scripted Cooperation														Task 1	0.384	0.29 0	0.158	0.288
						Scripted Cooperation														Task 2	-0.065	0.28 7		
						Scripted Cooperation														Completeness score initial	0.287	0.37 8		
						Scripted Cooperation														Completeness score transfer	-0.009	0.37 6		
						Scripted Cooperation														Communicati veness score initial	1.060	0.39 8	0.543	0.382
						Scripted Cooperation														Communicati veness score transfer	0.915	0.39 2		
						Scripted Cooperation														Completeness score initial	1.205	0.40 4		
						Scripted Cooperation														Completeness score transfer	0.129	0.37 7		
						Scripted Cooperation														Communicati veness score initial	0.530	0.38 2	0.522	0.382
						Scripted Cooperation														Communicati veness score transfer	0.308	0.37 8		
						Scripted Cooperation														Adjusted Means Score on a teacher- made test	-0.298	0.28 1		
						Scripted Cooperation														Adjusted Means First program	0.051	0.28 0	0.139	0.280
						Scripted Cooperation														Adjusted Means Second program	-0.173	0.28 0		
						Scripted Cooperation														Adjusted Means Score on a teacher- made test	-0.086	0.28 2		
						Scripted Cooperation														Adjusted Means First program	0.285	0.28 4	0.160	0.283



75. Roen, R., Van Pilsun, J., Harris, I., Rosenberg, E., Johnson, R., Law, C. & Rosenthal, L. (1983)	1	4	5	50 first-year medical students at the University of Minnesota. Pilot study in Biochemistry Department	Química	Cooperative Learning	5	2,4	1	2										Cooperative learning group, Ne=24; control group, Ne=25	Total course examination points	1,172	0,309
76. Şahin, A. (2010)	3	4	2	80 undergraduate Turkish language teacher education students who were attending Written Expression course classes at Atatürk University	Lengauje Escrito	Jigsaw II	2	6	1	2	1	1	1	2	1	1	3		Jigsaw II, Ne=42; control group, Ne=38	Gains Written Expression Achievement Test (WEAT)	0,901	0,235	



[illegible]





